

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 21:08:17 ; Search time 2.8 Seconds
(without alignments)
721.436 Million cell updates/sec

Title: US-09-980-381A-70

Perfect score: 110

Sequence: 1 AANARERRRHGLNHFADQLR 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	351	2 A56387	helix-loop-helix t
2	84	76.4	312	2 A40708	basic-helix-loop-h
3	81	73.6	337	2 I48682	NEX-1 - mouse
4	81	73.6	337	2 I57038	gene Dlx-3 protein
5	81	73.6	352	2 I51687	neurogenic differe
6	80	72.7	347	2 JC7999	basic helix-loop-h
7	80	72.7	357	2 JC4703	neurogenic differe
8	80	72.7	357	2 I49338	beta-cell E-box tr
9	80	72.7	381	2 A57059	hypothetical prote
10	78	70.9	192	2 T15764	helix-loop-helix t
11	77	70.0	367	2 JC6087	LIN-32 protein - C
12	73	66.4	71	2 T29378	neurogenic basic-h
13	73	66.4	381	2 G02668	KW8 protein - rat
14	73	66.4	381	2 JC4647	neuro D-related fa
15	73	66.4	383	2 JC4688	hypothetical prote
16	71	64.5	237	2 T15548	transcription fact
17	70	63.6	326	2 S71755	hypothetical prote
18	67	60.9	207	2 T20391	protein kinase C-b
19	65	59.1	334	2 G02409	hypothetical prote
20	63	57.3	258	2 T21959	hypothetical prote
21	62	56.4	170	2 T29971	basic helix-loop-h
22	62	56.4	360	2 A56086	Xtvi protein - Afr
23	59	53.6	166	2 A33637	Myogenic factor -
24	59	53.6	255	2 JN0624	transcription fact
25	57	51.8	198	2 A57717	gene achaete prote
26	56	50.9	201	2 S23324	Achaete-scute comp
27	56	50.9	201	2 A43731	twist protein homo
28	56	50.9	201	2 G01204	gene M-twist prote
29	56	50.9	206	2 I53066	

30 56 50.9 240 2 B43909 myogenic protein M
31 56 50.9 242 2 A34872 muscle-specific re
32 56 50.9 242 2 A42668 myogenic factor MR
33 56 50.9 242 2 S12385 myf-6 protein - hu
34 56 50.9 242 2 JC1233 muscle-specific re
35 56 50.9 260 2 A41123 myogenic transcrip
36 55 50.0 64 2 A48411 Myf5 homolog - chl
37 55 50.0 190 2 B43783 myogenesis protein
38 55 50.0 246 2 B43912 myogenic regulator
39 55 50.0 255 2 S16151 myogenic factor Xn
40 55 50.0 255 2 S22825 myf-5 protein - mc
41 55 50.0 255 2 S04105 myogenic factor 5
42 55 50.0 258 2 S41126 Myf5 protein - chl
43 55 50.0 289 2 A34783 myogenesis protein
44 55 50.0 297 2 A35874 myogenic factor 1
45 55 50.0 298 2 A32872 myogenic factor CM

ALIGNMENTS

RESULT 1

A56387

helix-loop-helix transcription factor MATH-1 - mouse

N;Alternate names: atonal homolog MATH-1

C;Species: Mus musculus (house mouse)

C;Date: 21-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 05-Nov-1999

C;Accession: A56387

R;Akazawa, C.; Ishibashi, M.; Shimizu, C.; Nakanishi, S.; Kageyama, R.

J. Biol. Chem. 270, 8730-8738, 1995

A;Title: A mammalian helix-loop-helix factor structurally related to the product of Dros

A;Reference number: A56387; MUID:95238366; PMID:7721778

A;Accession: A56387

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: DNA

A;Residues: 1-351 <AA>

A;Cross-references: GB:D43694; NID:g994770; PIDN:BAA07791.1; PID:d1008377; PID:g994771

C;Genetics:

A;Introns: #status absent

C;Keywords: transcription factor

Query Match 100.0%; Score 110; DB 2; Length 351;
Best Local Similarity 100.0%; Pred. No. 9.5e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRRHGLNHFADQLR 21

|||||

Db 160 AANARERRRHGLNHFADQLR 180

RESULT 2

A40708

basic-helix-loop-helix protein ato - fruit fly (Drosophila melanogaster)

C;Species: Drosophila melanogaster

C;Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000

C;Accession: A40708

R;Jarman, A.P.; Grau, Y.; Jan, L.Y.; Jan, Y.N.

Cell 73, 1307-1321, 1993

A;Title: atonal is a proneural gene that directs chordotonal organ formation in the Dros

A;Reference number: A40708; MUID:93313961; PMID:8324823

A;Accession: A40708

A;Status: preliminary

A;Molecule type: nucleic acid

A;Residues: 1-312 <JAR>

A;Cross-references: GB:I36646; NID:g551565; PIDN:AAA21879.1; PID:g551566

A;Note: sequence extracted from NCBI backbone (NCBIN:135094, NCBIP:135095)

C;Genetics:

A;Gene: FlyBase:ato

A;Cross-references: FlyBase:FBgn0010433

Query Match

Best Local Similarity 76.4%; Score 84; DB 2; Length 312;

Matches 81.0%; Pred. No. 1.1e-05;

```

Matches 17; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AANARERRMHGHNHAFDQLR 21
||||| ||| ||| |||
259 AANARERRMQLNQAFLRL 279

DBD

RESULT 3
I48682
NEX-1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 20-Jun-2000
C:Accession: I48682; S69346
R:R.Bartholoma, A.; Nave, K.A.
Mech. Dev. 48, 217-228, 1994
A:Title: NEX-1: a novel brain-specific helix-loop-helix protein with autoregulation and
A:Reference number: I48682; MUID:95200803; PMID:7545978
A:Accession: I48682
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-337 <RES>
A:Cross-references: EMBL:U29086; NID:g881961; PIDN:AAL4576.1; PID:g881962
R:R.Shimizu, C.; Akazawa, C.; Nakanishi, S.; Kageyama, R.
Eur. J. Biochem. 229, 239-248, 1995
A:Title: MATH-2, a mammalian helix-loop-helix factor structurally related to the product
A:Reference number: S69346; MUID:95262673; PMID:7744035
A:Accession: S69346
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-337 <SHI>
A:Cross-references: EMBL:D44480; NID:g994772; PIDN:BAA07923.1; PID:g994773
C:Genetics:
A:Gene: nex-1

Query Match 73.6%; Score 81; DB 2; Length 337;
Best Local Similarity 80.0%; Pred. No. 3.5e-05;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AANARERRMHGHNHAFDQLR 21
||||| ||| ||| |||
99 AANARERNMGINDALDNL 118

DBD

RESULT 4
I57038
gene Dlx-3 protein - mouse
C:Species: Mus sp. (mouse)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 28-Feb-1997
C:Accession: I57038
R:Robinson, G.W.; Mahon, K.A.
Mech. Dev. 48, 189-215, 1994
A:Title: Differential and overlapping expression domains of Dlx-2 and Dlx-3 suggest dist
A:Reference number: I57038; MUID:95200802; PMID:7893603
A:Accession: I57038
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-337 <RES>
A:Cross-references: GB:S75299; NID:g896199; PID:g896200
C:Genetics:
A:Gene: Dlx-3

Query Match 73.6%; Score 81; DB 2; Length 337;
Best Local Similarity 80.0%; Pred. No. 3.5e-05;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AANARERRMHGHNHAFDQLR 21
||||| ||| ||| |||
99 AANARERNMGINDALDNL 118

DBD

RESULT 5
I51687
neurogenic differentiation factor - African clawed frog

```

F:102-157/Region: helix-loop-helix #status predicted

Query Match 72.7%; Score 80; DB 2; Length 357;
Best Local Similarity 80.0%; Pred. No. 5.4e-05;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ANARERRMHGHLNHAFFDQLR 21
||||| ||||| ||||| |||||

DB 106 ANARERNMHGHLNHAFFDQLR 125

RESULT 8

I49338

neurogenic differentiation factor - mouse

C:Species: Mus musculus (house mouse)

C>Date: 02-Jul-1996 #sequence_revision 16-Aug-1996 #text_change 05-Nov-1999

C:Accession: I49338; I49345

R:Lee, J.E.; Hollenberg, S.M.; Snider, L.; Turner, D.L.; Lipnick, N.; Weintraub, H.

Science 268, 836-844, 1995

A:Title: Conversion of Xenopus ectoderm into neurons by NeuroD, a basic helix-loop-helix

A:Reference number: A56481; MUID:95273957; PMID:7754368

A:Accession: I49338

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 73-357 <RE2>

A:Cross-references: EMBL:U28068; NID:G854738; PIDN:AA52203.1; PID:G854739

A:Accession: I49345

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-121 <RE3>

A:Cross-references: EMBL:U28888; NID:G854742; PIDN:AA52204.1; PID:G854743

C:Genetics:

A:Gene: neuroD

Query Match

Best Local Similarity 80.0%; Pred. No. 5.4e-05;

Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ANARERRMHGHLNHAFFDQLR 21

||||| ||||| ||||| |||||

DB 106 ANARERNMHGHLNHAFFDQLR 125

RESULT 9

A57059

beta-cell E-box transcription activator 2 - hamster (fragment)

C:Species: Cricetinae gen. sp. (hamster)

C>Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 08-Oct-1999

C:Accession: A57059

R:Naya, F.J.; Stellrecht, C.M.M.; Tsai, M.J.

Genes Dev. 9, 1009-1019, 1995

A:Title: Tissue-specific regulation of the insulin gene by a novel basic helix-loop-helix

A:Reference number: A57059; MUID:95293222; PMID:7774807

A:Accession: A57059

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-381 <NAY>

A:Cross-references: GB:U24679; NID:G777418; PIDN:AAA86518.1; PID:G777419

Query Match

Best Local Similarity 80.0%; Pred. No. 5.8e-05;

Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ANARERRMHGHLNHAFFDQLR 21

||||| ||||| ||||| |||||

DB 131 ANARERNMHGHLNHAFFDQLR 150

RESULT 10

T15764

hypothetical protein C34E10.7 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T15764

R:Kirsten, J.

submitted to the EMBL Data Library, June 1994

A:Description: The sequence of C. elegans cosmid C34E10.

A:Reference number: Z18399

A:Accession: T15764

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-192 <KIR>

A:Cross-references: EMBL:U10402; NID:G500723; PID:G500730; PIDN:AAA19069.1; CESP:C34E10.

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:C34E10.7

A:Introns: 65/2; 107/3

Query Match

Best Local Similarity 75.0%; Pred. No. 5.8e-05;

Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ANARERRMHGHLNHAFFDQLR 21

||||| ||||| ||||| |||||

DB 24 ANGRERARMHGLNNAIDMLR 43

RESULT 11

JC6087

helix-loop-helix transcription factor, BETA3 - hamster

C:Species: Cricetinae gen. sp. (hamster)

C>Date: 28-May-1997 #sequence_revision 18-Jul-1997 #text_change 05-Nov-1999

C:Accession: JC6087

R:Pepton, M.; Stellrecht, C.M.M.; Naya, F.J.; Huang, H.P.; Samora, P.J.; Tsai, M.J.

Mol. Cell. Biol. 16, 626-633, 1996

A:Title: BETA3, a novel helix-loop-helix protein, can act as a negative regulator of BBT

A:Reference number: JC6087; MUID:96140430; PMID:8552091

A:Accession: JC6087

A:Molecule type: mRNA

A:Residues: 1-367 <PEY>

A:Cross-references: GB:S80870; NID:G1911496; PIDN:AAB50691.1; PID:G1911497

A:Experimental source: insulin tumor cell

C:Comment: This factor belongs to the tissue-specific class B basic helix-loop-helix fam

ssion.

C:Genetics:

A:Gene: beta3

C:Keywords: repressor; transcription factor

Query Match

Best Local Similarity 78.9%; Pred. No. 0.00016;

Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 NARERRRMHGLNHAFFDQLR 21

||||| ||||| ||||| |||||

DB 234 NARERRRMHGLNHAFFDQLR 252

RESULT 12

T29378

Lin-32 protein - Caenorhabditis elegans

C:Alternate names: hlx transcription factor

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000

C:Accession: T29378; S52415

R:Miller, N.; Langston, Y.

submitted to the EMBL Data Library, February 1996

A:Description: The sequence of C. elegans cosmid T14F9.

A:Reference number: Z20615

A:Accession: T29378

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-71 <MIL>

A:Cross-references: EMBL:U50199; PIDN:AAA91264.1; CESP:lin-32

A:Experimental source: strain Bristol N2

R:Zhao, C.; Emmons, S.W.

Nature 373, 74-78, 1995

A:Title: A transcription factor controlling development of peripheral sense organs in C.
 A:Reference number: S52415; MUID:95098129; PMID:7800042
 A:Accession: S52415

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-71 <ZHA>

A:Cross-references: EMBL:U15418; NID:g560021; PIDN:AAA67360.1; PID:g560022

C:Genetics:

A:Gene: CESP:lin-32

A:Introns: 23/3

Query Match 66.4%; Score 73; DB 2; Length 71;
 Best Local Similarity 71.4%; Pred. No. 0.00013;
 Matches 15; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ANARERRRMHGLNHFADQLR 21

Db 5 AANERERRMTNLNVAYDEL 25

RESULT 13

G02668

C:Species: neurogenic basic-helix-loop-helix (bHLH) protein - human

C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 17-Jul-1998

C:Accession: G02668

R:Tapscott, S.J.; Tamimi, R.T.; McCormick, B.M.

A:Reference number: H01573

A:Accession: G02668

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-381 <TAP>

A:Cross-references: EMBL:U58681; NID:g1477748; PID:g1477749

C:Genetics:

A:Gene: NeuroD2

Query Match 66.4%; Score 73; DB 2; Length 381;
 Best Local Similarity 75.0%; Pred. No. 0.00074;
 Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 ANARERRRMHGLNHFADQLR 21

Db 126 ANARERNRMDLNAALDNL 145

RESULT 14

JC4647

N:KW8 protein - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 20-Jun-2000

C:Accession: JC4647

R:Kume, H.; Maruyama, K.; Tomita, T.; Iwatsubo, T.; Saïdo, T.C.; Obata, K.

A:Title: Molecular cloning of a novel basic helix-loop-helix protein from the rat brain.

A:Reference number: JC4647; MUID:96193685; PMID:8605021

A:Accession: JC4647

A:Molecule type: DNA

A:Residues: 1-381 <KUM>

A:Cross-references: DDBJ:D82868; NID:g1166397; PIDN:BA11615.1; PID:g1166398

A:Experimental source: brain

C:Comment: This protein is involved in synaptic plasticity, and has a role specific to a

-loop-helix domain.

C:Keywords: brain

F:122-134/Region: basic

Query Match 66.4%; Score 73; DB 2; Length 381;
 Best Local Similarity 75.0%; Pred. No. 0.00074;
 Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 ANARERRRMHGLNHFADQLR 21

Db 126 ANARERNRMDLNAALDNL 145

Db 126 ANARERNRMDLNAALDNL 145

RESULT 15

JC4688

C:Species: neuro D-related factor - mouse

C:Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 20-Jun-2000

C:Accession: JC4688

R:Yasunami, M.; Suzuki, K.; Maruyama, H.; Kawakami, H.; Nagai, Y.; Hagiwara, M.; Ohkubo,

Biochem. Biophys. Res. Commun. 220, 754-758, 1996

A:Title: Molecular cloning and characterization of a cDNA encoding a novel basic helix-

A:Reference number: JC4688; MUID:96183021; PMID:8607837

A:Accession: JC4688

A:Molecule type: mRNA

A:Residues: 1-383 <YAS>

A:Cross-references: DDBJ:D83507; NID:g1304163; PIDN:BAA11931.1; PID:g1304164

A:Experimental source: embryo

C:Comment: This factor plays distinct roles in neural development and plasticity as a

F:36-43/Region: proline-rich

F:83-92/Region: glutamic acid-rich

F:213-237/Region: alanine/glycine-rich

Query Match 66.4%; Score 73; DB 2; Length 383;
 Best Local Similarity 75.0%; Pred. No. 0.00074;
 Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 ANARERRRMHGLNHFADQLR 21

Db 127 ANARERNRMDLNAALDNL 146

Search completed: September 21, 2004, 21:25:00

Job time : 2.8 secs

Result No.	Score	Query		ID	Description
		Match	Length		
1	110	100.0	351	1 ATH1_MOUSE	P49885 mus musculus
2	110	100.0	354	1 ATH1_HUMAN	Q92858 homo sapien
3	84	76.4	198	1 AMOS_DROME	Q970a7 drosophila
4	84	76.4	312	1 ATO_DROME	P48987 drosophila
5	81	73.6	337	1 NDF6_HUMAN	Q96nks homo sapien
6	81	73.6	337	1 NDF6_MOUSE	P48986 mus musculus
7	81	73.6	352	1 NDF1_XENLA	Q91616 xenopus lae
8	80	72.7	330	1 NDF4_MOUSE	O03105 mus musculus
9	80	72.7	330	1 NDFM_CHICK	P79766 gallus gall
10	80	72.7	331	1 NDF4_HUMAN	Q9hd90 homo sapien
11	80	72.7	355	1 NDF1_MESAU	Q60430 mesocricetu
12	80	72.7	356	1 NDF1_HUMAN	Q13562 homo sapien
13	80	72.7	357	1 NDF1_CHICK	P79765 gallus gall
14	80	72.7	357	1 NDF1_MOUSE	Q60867 mus musculus
15	80	72.7	357	1 NDF1_RAT	P64289 rattus norv
16	78	70.9	192	1 YLB7_CARBL	P46581 caenorhabdi
17	77	70.0	367	1 BET3_MESAU	O90929 mesocricetu
18	76	69.1	315	1 NDF3_XENLA	P79920 xenopus lae
19	73	66.4	71	1 LI32_CARBL	Q10574 caenorhabdi
20	73	66.4	382	1 NDF2_HUMAN	Q15784 homo sapien
21	73	66.4	382	1 NDF2_RAT	Q63689 rattus norv
22	73	66.4	383	1 NDF2_MOUSE	Q62414 mus musculus
23	68	61.8	208	1 NGN1_BRABE	Q42606 brachydanio
24	68	61.8	214	1 NGN3_MOUSE	P70661 mus musculus
25	67	60.9	214	1 NGN3_HUMAN	Q9y422 homo sapien
26	67	60.9	237	1 NGN1_HUMAN	Q92886 homo sapien
27	67	60.9	244	1 NGN1_MOUSE	P70660 mus musculus
28	67	60.9	244	1 NGN1_RAT	P70595 rattus norv
29	67	60.9	250	1 NGN2_HUMAN	Q9h2a3 homo sapien
30	67	60.9	263	1 NGN2_MOUSE	P70447 mus musculus
31	66	60.0	398	1 TAP_DROME	O16867 drosophila
32	65	59.1	298	1 OLG2_CHICK	Q90xb3 gallus gall
33	65	59.1	323	1 OLG2_HUMAN	Q13516 homo sapien

CC epithelial cells of the gastrointestinal tract.
 CC -!- DEVELOPMENTAL STAGE: First detected in the cranial ganglions and
 CC the dorsal part of the central nervous system on embryonic day 9.5
 CC (E9.5). From E10.5 onward, prominent expression of MATH-1
 CC continues in the dorsal part of the central nervous system but
 CC becomes restricted to the external granular layer of the
 CC cerebellum by e18 and is undetectable in the adult nervous system.
 CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; D43694; BAA07791.1; -;
 CC EMBL; BC010820; AAH10820.1; -;
 CC EMBL; BC051256; AAH51256.1; -;
 CC PIR; A56387; A56387.
 CC TRANSFAC; T01668; -;
 CC MGD; MGI:104654; Atoh1.
 CC GO; GO:0007420; P:brain development; IMP.
 CC InterPro; IPR001092; HLH_basic.
 CC Pfam; PF00010; HLH; 1.
 CC SMART; SM00353; HLH; 1.
 CC PROSITE; PS50888; HLH; 1.
 CC Transcription regulation; Activator; DNA-binding; Nuclear protein.
 CC DNA_BIND 157 168 BASIC DOMAIN.
 CC FT DOMAIN 169 209 HELIX-LOOP-HELIX MOTIF.
 CC FT DOMAIN 221 224 POLY-PRO.
 CC SEQUENCE 351 AA; 37854 MW; 596E03DF23C3BF96 CRC64;
 CC
 CC Query Match 100.0%; Score 110; DB 1; Length 351;
 CC Best Local Similarity 100.0%; Pred. No. 2.5e-10;
 CC Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 1 AANARERRRMHGLNHAFDQLR 21
 CC |||||
 CC DB 160 AANARERRRMHGLNHAFDQLR 180
 CC
 CC RESULT 2
 CC ATTH_HUMAN STANDARD; PRT; 354 AA.
 CC ID ATTH_HUMAN
 CC AC Q92858;
 CC DT 15-DEC-1998 (Rel. 37, Created)
 CC DT 15-DEC-1998 (Rel. 37, Last sequence update)
 CC DT 15-MAR-2004 (Rel. 43, Last annotation update)
 CC DE Atonal protein homolog 1 (Helix-loop-helix protein bATH-1).
 CC GN ATOH1 OR ATH1.
 CC OS Homo sapiens (Human).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CC OX NCBI_TaxID=9606;
 CC RN [1]
 CC SEQUENCE FROM N.A.
 CC RP MEDLINE=97026280; PubMed=8872459;
 CC RA Ben-Arie N., McCall A.E., Berkman S., Eichele G., Bellen H.J.,
 CC Zoghbi H.Y.
 CC RT "Evolutionary conservation of sequence and expression of the bHLH
 CC protein Atonal suggests a conserved role in neurogenesis.";
 CC RL Hum. Mol. Genet. 5:1207-1216(1996).
 CC -!- FUNCTION: Activates E box-dependent transcription in collaboration
 CC with E47, but the activity is completely antagonized by the
 CC negative regulator of neurogenesis HES1. May play a role in the
 CC differentiation of subsets of neural cells by activating E box-
 CC dependent transcription (By similarity).
 CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
 CC bHLH protein.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U61148; AAB41305.1; -;
 CC TRANSFAC; T04544; -;
 CC Genew; HGNC:797; ATOH1.
 CC MIM; 601461; -;
 CC GO; GO:0003700; P:transcription factor activity; TAS.
 CC GO; GO:0007417; P:central nervous system development; TAS.
 CC GO; GO:0006366; P:transcription from Pol II promoter; TAS.
 CC InterPro; IPR001092; HLH_basic.
 CC Pfam; PF00010; HLH; 1.
 CC SMART; SM00353; HLH; 1.
 CC PROSITE; PS50888; HLH; 1.
 CC Transcription regulation; Activator; DNA-binding; Nuclear protein.
 CC DOMAIN 29 38 POLY-PRO.
 CC FT DNA_BIND 160 171 BASIC DOMAIN.
 CC FT DOMAIN 172 212 HELIX-LOOP-HELIX MOTIF.
 CC FT DOMAIN 224 228 POLY-PRO.
 CC SEQUENCE 354 AA; 38160 MW; AB12F1E917A00A8D CRC64;
 CC
 CC Query Match 100.0%; Score 110; DB 1; Length 354;
 CC Best Local Similarity 100.0%; Pred. No. 2.5e-10;
 CC Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 1 AANARERRRMHGLNHAFDQLR 21
 CC |||||
 CC DB 163 AANARERRRMHGLNHAFDQLR 183
 CC
 CC RESULT 3
 CC AMOS DROME STANDARD; PRT; 198 AA.
 CC ID AMOS DROME
 CC AC Q9Y0A7; Q9YV76;
 CC DT 28-FEB-2003 (Rel. 41, Created)
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
 CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
 CC DE Basic helix-loop-helix transcription factor Amos (Reduced olfactory
 CC DE organs protein) (Rough eye protein) (Absent MD neurons and olfactory
 CC DE sensilla protein) (Amos protein).
 CC GN AMOS OR ROI OR ROLO OR CG10393.
 CC OS Drosophila melanogaster (Fruit fly).
 CC OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC OC Ephydroidea; Drosophilidae; Drosophila.
 CC OX NCBI_TaxID=7227;
 CC RN [1]
 CC SEQUENCE FROM N.A. AND FUNCTION.
 CC RP STRAIN=Oregon-R;
 CC RX MEDLINE=20170246; PubMed=10707973;
 CC RA Goulding S.E., zur Lage P., Jarman A.P.;
 CC RT "Amos, a proneural gene for Drosophila olfactory sense organs that is
 CC RT regulated by lozenge";
 CC RL Neuron 25:69-78(2000).
 CC RN [2]
 CC SEQUENCE FROM N.A.
 CC RP STRAIN=Berkely;
 CC RX MEDLINE=20196006; PubMed=107311132;
 CC RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 CC RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 CC RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 CC RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 CC RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champe M., Pfeiffer B.D.,
 CC RA Wank K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 CC RA Abriel J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 CC RA Balwle R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 CC RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Deicher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulyov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195 (2000).
RN [3]
RP FUNCTION.
RX MEDLINE=20170245; PubMed=10707972;
RA Huang M.L., Hsu C.H., Chien C.T.;
RT "The proneural gene *amos* promotes multiple dendritic neuron formation
in the *Drosophila* peripheral nervous system.";
RL Neuron 25:57-67 (2000).
CC -!- FUNCTION: Transcription factor involved in early neurogenesis.
CC Promotes multiple dendritic (MD) neuron formation. Required for
CC olfactory sensilla.
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein. Interacts with Daughterless.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- DEVELOPMENTAL STAGE: During embryonic development, *amos* is
CC expressed in patches of ectodermal cells, and the expression is
CC quickly restricted to sensory organ precursors.
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch.
CC -----
CC EMBL; AF166113; AAD45410.1; -;
CC DR EMBL; AE003659; AAF53678.1; -;
CC DR HSPB; P25912; 1HLO.
CC DR FlyBase; FBgn0003270; *amos*.
CC DR InterPro; IPR001092; HLH_basic.
CC DR Pfam; PF00010; HLH; 1.
CC DR SMART; SM00353; HLH; 1.
CC DR PROSITE; PS50888; HLH; 1.
KW Nuclear protein; Transcription regulation; Developmental protein;
KW Neurogenesis.
FT DOMAIN 111 128 SRR-RICH.
FT DNA_BIND 138 150 BASIC DOMAIN.
FT DOMAIN 151 191 HELIX-LOOP-HELIX MOTIF.
FT CONFLICT 25 25 S -> G (IN REF. 1).
SQ SEQUENCE 198 AA; 22569 MW; F1D22AF2BA1C670B CRC64;

Query Match 76.4%; Score 84; DB 1; Length 198;
Best Local Similarity 81.0%; Pred. No. 2e-06;

Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 AANARERRRMHGLNFAFDQLR 21
DB 142 AANARERRRMNSINDAFDKLR 162
RESULT 4
ATO DROME STANDARD; PRT; 312 AA.
AC ATO DROME Q9VHUH;
DT 01-FEB-1996 (Rel. 33, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Atonal protein.
GN ATO OR CG7508.
OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; *Drosophila*.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
RC STRAIN=Oregon-R;
RX MEDLINE=93313961; PubMed=8324823;
RA Jarman A.P., Grau Y., Jan L.Y., Jan Y.N.;
RT "Atonal is a proneural gene that directs chordotonal organ formation
in the *Drosophila* peripheral nervous system.";
RL Cell 73:1307-1321 (1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle A.C., Baxter E.G., Helt G., Nelson C.R., Miklos G.D.G.,
RA April J.F., Agayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Deicher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulyov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195 (2000).
RN [3]
RP FUNCTION.

```

RX MEDLINE=94255014; PubMed=8196767;
RA Jarman A.P., Grell E.H., Ackerman L., Jan L.Y., Jan Y.N.;
RT "Atonal is the proneural gene for Drosophila photoreceptors.";
RL Nature 369:398-400(1994).
CC -!- FUNCTION: Developmental protein involved in neurogenesis. Required
CC for the formation of chordotonal organs and photoreceptors. Seems
CC to bind to E boxes. Specifically required for the photoreceptor R8
CC selection.
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein. Forms a heterodimer with Daughterless.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Proneural clusters and sense organ precursors
CC of the chordotonal organs, optic furrow of the eye-antennal disk
CC and developing brain lobe.
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L36646; AAA21879.1; -.
CC EMBL; AE003678; AAF54209.1; -.
CC PIR; A40708; A40708.
CC FlyBase; FBgn010433; ato.
CC GO; GO:0007420; P:brain development; IMP.
CC GO; GO:0007173; P:EGF receptor signaling pathway; IGI.
CC GO; GO:0007605; P:hearing; IMP.
CC GO; GO:0007438; P:neocortex development; IMP.
CC GO; GO:0007422; P:peripheral nervous system development; NAS.
CC GO; GO:0045464; P:R8 cell fate specification; NAS.
CC GO; GO:0007224; P:smoothed signaling pathway; IGI.
CC InterPro; IPR001092; HLH_basic.
CC Pfam; PF00010; HLH; 1.
CC SMART; SM00353; HLH; 1.
CC PROSITE; PS00888; HLH; 1.
KW Neurogenesis; Differentiation; Developmental protein; Nuclear protein;
KW Transcription regulation; DNA-binding.
FT DNA BIND 255 267 BASIC DOMAIN.
FT DOMAIN 268 308 HELIX-LOOP-HELIX MOTIF.
FT CONFLICT 149 149 G -> A (IN REF. 1).
FT CONFLICT 149 149
SQ SEQUENCE 312 AA; 34116 MW; 069479287438F456 CRC64;

Query Match 76.4%; Score 84; DB 1; Length 312;
Best Local Similarity 81.0%; Pred. No. 3.3e-06;
Matches 17; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AANARERRRMHGLNHFADQLR 21
Db 259 AANARERRRMQNQAFLRLR 279
|||||
|||||

RESULT 5
NDP6 HUMAN STANDARD; PRT; 337 AA.
AC Q96NR8; Q9H3H6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Neurogenic differentiation factor 6 (NeuroD6) (Myo51 protein).
GN NEUROD6
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RA Mao Y.M., Xie Y., Zheng Z.H.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.

```

```

RN SEQUENCE FROM N.A.
RP TISSUE=Brain;
RA Tashiro H., Yanagaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Koniyama M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,
RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuna M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,
RA Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K.,
RA Masuno Y., Nagai K., Isogai T.;
RA "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Activates E box-dependent transcription in collaboration
CC with E47. May be a trans-acting factor involved in the development
CC and maintenance of the mammalian nervous system. Transactivates
CC the promoter of its own gene (By similarity).
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC -!- CAUTION: Ref.1 sequence differs from that shown due to frameshifts
CC in positions 300 and 307.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF063609; AAC43167.1; ALT_FRAME.
CC EMBL; AK055238; BAB70885.1; -.
CC Genew; HGNC:13804; NEUROD6.
CC InterPro; IPR001092; HLH_basic.
CC Pfam; PF00010; HLH; 1.
CC SMART; SM00353; HLH; 1.
CC PROSITE; PS00888; HLH; 1.
KW Transcription regulation; Activator; DNA-binding; Nuclear protein.
FT DOMAIN 54 63 POLY-GLU.
FT DOMAIN 80 86 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 95 106 BASIC DOMAIN.
FT DOMAIN 107 147 HELIX-LOOP-HELIX MOTIF.
FT CONFLICT 107 147
SQ SEQUENCE 337 AA; 38705 MW; 6B0F4127AC0F809E CRC64;

Query Match 73.6%; Score 81; DB 1; Length 337;
Best Local Similarity 80.8%; Pred. No. 1.1e-05;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AANARERRRMHGLNHFADQLR 21
Db 99 AANARERRRMGLNDALDNLR 118
|||||
|||||

RESULT 6
NDP6 MOUSE STANDARD; PRT; 337 AA.
AC P48966;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neurogenic differentiation factor 6 (NeuroD6)
DE 2) (Helix-loop-helix protein MATH-2) (MATH2) (NEX-1 protein).
GN NEUROD6 OR ATOH2 OR ATH2 OR NEX1.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/J;
RA
RL MEDLINE=95262673; PubMed=7744035;

```


RA Shimizu C., Akazawa C., Nakanishi S., Kageyama R.;
 RT "MATH-2, a mammalian helix-loop-helix factor structurally related to
 RT the product of Drosophila proneural gene atonal, is specifically
 RT expressed in the nervous system.";
 RL Eur. J. Biochem. 229:239-248(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=95200803; PubMed=7545978;
 RA Bartholomae A., Nave K.-A.;
 RT "NEX-1: a novel brain-specific helix-loop-helix protein with
 RT autoregulation and sustained expression in mature cortical neurons.";
 RL Mech. Dev. 48:217-228(1994).
 CC -!- FUNCTION: Activates E box-dependent transcription in collaboration
 CC with E47. May be a trans-acting factor involved in the development
 CC and maintenance of the mammalian nervous system. Transactivates
 CC the promoter of its own gene.
 CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
 CC bHLH protein.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- TISSUE SPECIFICITY: Specific to the nervous system of both embryos
 CC and adults. Highest levels in the cortical plate of the cerebrum.
 CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL; D44480; BAA07923.1; -;
 DR EMBL; U29086; AAC14576.1; -;
 DR PIR; I48682; I48682.
 DR PIR; I57038; I57038.
 DR MGD; MGI:106593; Neurod6.
 DR InterPro; IPR001092; HLH_basic.
 DR Pfam; PF00010; HLH; 1.
 DR SMART; SM00353; HLH; 1.
 DR PROSITE; PS50888; HLH; 1.
 KW Transcription regulation; Activator; DNA-binding; Nuclear protein.
 FT DOMAIN 54 63 POLY-GLU.
 FT DOMAIN 80 86 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DNA_BIND 95 106 BASIC DOMAIN.
 FT DOMAIN 107 147 HELIX-LOOP-HELIX MOTIF.
 SQ SEQUENCE 337 AA; 38644 MW; 35C18ACD8EE1EFBA CRC64;
 Query Match 73.6%; Score 81; DB 1; Length 337;
 Best Local Similarity 80.0%; Pred. No. 1.1e-05;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 2 ANARERRMHGHLNHAFDQLR 21
 DB 99 ANARERNRHHGLNDALNLR 118
 RESULT 7
 NDFF1_XENLA
 ID NDFF1_XENLA STANDARD; PRT; 352 AA.
 AC Q91616;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Neurogenic differentiation factor 1 (Neurod1).
 GN NEUROD1 OR NEUROD.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 CC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=95273957; PubMed=7754368;
 RA Lee J.E., Hollenberg S.M., Snider L., Turner D.L., Lipnick N.,
 RA Weintraub H.;
 RT "Conversion of Xenopus ectoderm into neurons by NeuroD, a basic
 RT helix-loop-helix protein.";
 RL Science 268:836-844(1995).
 CC -!- FUNCTION: Acts as a differentiation factor during neurogenesis.
 CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
 CC bHLH protein.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -!- TISSUE SPECIFICITY: Expressed in differentiating neurons of
 CC both the central and peripheral nervous systems.
 CC -!- DEVELOPMENTAL STAGE: Expressed during embryonic development.
 CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL; U28067; AAC59675.1; -;
 DR PIR; I51687; I51687.
 DR InterPro; IPR001092; HLH_basic.
 DR Pfam; PF00010; HLH; 1.
 DR SMART; SM00353; HLH; 1.
 DR PROSITE; PS50888; HLH; 1.
 KW DNA-binding; Nuclear protein; Transcription regulation; Activator;
 KW Neurogenesis; Developmental protein; Differentiation.
 FT DOMAIN 59 78 GLU-RICH (ACIDIC).
 FT DOMAIN 88 94 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DNA_BIND 103 114 BASIC DOMAIN.
 FT DOMAIN 115 155 HELIX-LOOP-HELIX MOTIF.
 SQ SEQUENCE 352 AA; 39662 MW; 226298DB3D48233E CRC64;
 Query Match 73.6%; Score 81; DB 1; Length 352;
 Best Local Similarity 80.0%; Pred. No. 1.1e-05;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 2 ANARERRMHGHLNHAFDQLR 21
 DB 107 ANARERNRHHGLNDALNLR 126
 RESULT 8
 NDFF4_MOUSE
 ID NDFF4_MOUSE STANDARD; PRT; 330 AA.
 AC O09105;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Neurogenic differentiation factor 4 (NeuroD4) (Atonal protein homolog
 DE 3) (Helix-loop-helix protein MATH-3) (MATH3).
 GN NEUROD4 OR ATOH3 OR ATH3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Retina;
 RX MEDLINE=97180939; PubMed=9029157;
 RA Takebayashi K., Takahashi S., Yokota C., Tsuda H., Nakanishi S.,
 RA Asashima M., Kageyama R.;
 RT "Conversion of ectoderm into a neural fate by ATH-3, a vertebrate
 RT basic helix-loop-helix gene homologous to Drosophila proneural gene
 RT atonal.";
 RL EMBO J. 16:384-395(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/J;

```
RX MEDLINE=98165816; PubMed=9497361;
RA Tsuda H., Takebayashi K., Nakanishi S., Kageyama R.;
RT "Structure and promoter analysis of Math3 gene, a mouse homolog of
RT Drosophila proneural gene atonal. Neural-specific expression by dual
RL J. Biol. Chem. 273:6327-6333 (1998).
RN [3]
RC SEQUENCE FROM N.A.
RP TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Heiton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RL human and mouse cDNA sequences."
RX Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Appears to mediate neuronal differentiation.
CC -1- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- TISSUE SPECIFICITY: Expressed in the developing nervous system,
CC with high levels of expression in the brain, retina and cranial
CC ganglions.
CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D85845; BAA12880.1; -.
CC EMBL; AF036257; AAC15969.1; -.
CC EMBL; BC054391; AAH54391.1; -.
CC MGD; MGI:108055; Neurod4.
CC GO; GO:0001654; P:eye morphogenesis; IMP.
CC GO; GO:0007400; P:neuroblast cell fate determination; IMP.
CC InterPro; IPR001092; HLH_basic.
CC Pfam; PF00010; HLH; 1.
CC SMART; SM00353; HLH; 1.
CC PROSITE; PS50888; HLH; 1.
CC DNA-binding; Nuclear protein; Transcription regulation; Activator;
CC Neurogenesis; Developmental protein; Differentiation.
CC FT DOMAIN 49 65 ASP/GLU-RICH (ACIDIC).
CC FT DOMAIN 55 61 POLY-LYS.
CC FT DOMAIN 73 76 POLY-LYS.
CC FT DOMAIN 73 79 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
CC PROSITE; PS50888; HLH; 1.
CC DNA-binding; Nuclear protein; Transcription regulation; Activator;
CC Neurogenesis; Developmental protein; Differentiation.
CC FT DOMAIN 49 65 ASP/GLU-RICH (ACIDIC).
CC FT DOMAIN 88 99 BASIC DOMAIN.
CC FT DOMAIN 100 140 HELIX-LOOP-HELIX MOTIF.
CC FT DOMAIN 162 183 LEUCINE-ZIPPER (POTENTIAL).
CC SQ SEQUENCE 330 AA; 37133 MW; P8EC228F4EE0FF88 CRC64;
Query Match 72.7%; Score 80; DB 1; Length 330;
Best Local Similarity 80.0%; Pred. No. 1.6e-05;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 ANARERRRMEGLNHAFFDLR 21
Db 92 ANARERTMEGLNDALDNL 111
RESULT 10
```

```
RESULT 9
ID NDFM_CHICK STANDARD; PRT; 330 AA.
AC P79756;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Neurogenic differentiation factor Neurom.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White leghorn; TISSUE=Retina;
RX MEDLINE=97454246; PubMed=9310321;
RA Roztocil T., Matter-Sadzinski L., Alliod C., Ballivet M.,
RA Matter J.M.;
RT "Neurom, a neural helix-loop-helix transcription factor, defines a
RT new transition stage in neurogenesis."
RL Development 124:3263-3272(1997).
CC -1- FUNCTION: Acts as a differentiation factor during neurogenesis.
CC Transcriptional activator. Binds DNA on E-box consensus sequence
CC 5'-CANNTG-3'.
CC -1- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- TISSUE SPECIFICITY: Developing nervous system.
CC -1- DEVELOPMENTAL STAGE: Expression in the developing nervous system
CC is transient and restricted to cells lining the ventricular zone
CC that have ceased proliferation but have not yet begun to migrate
CC into the outer layers. In retina, neurom is also transiently
CC expressed in cells as they withdraw from the mitotic cycle, but
CC persists in horizontal and bipolar neurons until full
CC differentiation. In the peripheral nervous system, its expression
CC closely follows cell proliferation.
CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y09597; CAA70785.1; -.
CC InterPro; IPR001092; HLH_basic.
CC Pfam; PF00010; HLH; 1.
CC SMART; SM00353; HLH; 1.
CC PROSITE; PS50888; HLH; 1.
CC DNA-binding; Nuclear protein; Transcription regulation; Activator;
CC Neurogenesis; Developmental protein; Differentiation.
CC FT DOMAIN 49 65 GLU-RICH (ACIDIC).
CC FT DOMAIN 55 61 POLY-LYS.
CC FT DOMAIN 73 76 POLY-LYS.
CC FT DOMAIN 73 79 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
CC PROSITE; PS50888; HLH; 1.
CC DNA BIND 88 99 BASIC DOMAIN.
CC FT DOMAIN 100 140 HELIX-LOOP-HELIX MOTIF.
CC FT DOMAIN 162 183 LEUCINE-ZIPPER (POTENTIAL).
CC SQ SEQUENCE 330 AA; 36707 MW; 96492F06A3C07B41 CRC64;
Query Match 72.7%; Score 80; DB 1; Length 330;
Best Local Similarity 80.0%; Pred. No. 1.6e-05;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 ANARERRRMEGLNHAFFDLR 21
Db 92 ANARERTMEGLNDALDNL 111
RESULT 10
```

```

NDFF4 HUMAN
ID NDF1 HUMAN STANDARD; PRT; 331 AA.
AC Q9HD90;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neurogenic differentiation factor 4 (Neurod4).
GN NEUROD4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANT THR-68.
RX MEDLINE=20527902; PubMed=11078465;
RA Horikawa Y., Horikawa Y., Cox N.J., Iwasaki N., Ogata M., Iwamoto Y.,
RA Schwitzgebel V., German M.S., Bell G.I.;
RT "beta-cell transcription factors and diabetes: no evidence for
RT diabetes-associated mutations in the gene encoding the basic
RT helix-loop-helix transcription factor neurogenic differentiation 4
RT (NEUROD4) in Japanese patients with MODY.";
RL Diabetes 49:1955-1957 (2000).
CC -!- FUNCTION: Appears to mediate neuronal differentiation (By
CC similarity).
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF203901; AAF99097.1; -.
CC Genew; HGNC:13802; NEUROD4.
CC InterPro; IPR001092; HLH_basic.
CC Pfam; PF00010; HLH; 1.
CC SMART; SM00353; HLH; 1.
CC PROSITE; PS50888; HLH; 1.
CC DNA-binding; Nuclear protein; Transcription regulation; Activator;
CC Neurogenesis; Developmental protein; Differentiation; Polymorphism.
CC DOMAIN 49 65
CC DNA BIND 88 99 BASIC DOMAIN.
CC DOMAIN 100 140 HELIX-LOOP-HELIX MOTIF.
CC VARIANT 68 68 K -> T.
CC /FTid=VAR_012979.
CC SEQUENCE 331 AA; 36982 MW; 56327EF2780ABF2B CRC64;

Query Match 72.7%; Score 80; DB 1; Length 331;
Best Local Similarity 80.0%; Pred. No. 1.6e-05;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ANARERRMHGHNHAFDQLR 21
| | | | | | | | | | | | | | | | | |
Db 92 ANARERTMHGNDALDNLRL 111

RESULT 11
NDF1 MESAU
ID NDF1 MESAU STANDARD; PRT; 355 AA.
AC Q60430;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neurogenic differentiation factor 1 (Neurod1) (Beta-cell E-box trans-
DE activator 2) (BETA2).
GN NEUROD1 OR NEUROD.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95293222; PubMed=7774807;
RA Naya F.J., Stellrecht C.M.M., Tsai M.-J.;
RA "Tissue-specific regulation of the insulin gene by a novel basic
RT helix-loop-helix transcription factor.";
RL Genes Dev. 9:1009-1019 (1995).
CC -!- FUNCTION: Acts as a differentiation factor during neurogenesis.
CC Transcriptional activator. Binds to the insulin gene E-box.
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein. Heterodimer with E47.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- TISSUE SPECIFICITY: Most abundant in pancreatic alpha- and beta-
CC cells, less in brain and intestine.
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U24679; AAA86518.1; ALT_INIT.
CC InterPro; IPR001092; HLH_basic.
CC Pfam; PF00010; HLH; 1.
CC SMART; SM00353; HLH; 1.
CC PROSITE; PS50888; HLH; 1.
CC DNA-binding; Nuclear protein; Transcription regulation; Activator;
CC Neurogenesis; Developmental protein; Differentiation.
CC DOMAIN 58 77 GLU-RICH (ACIDIC).
CC DOMAIN 86 92 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
CC DNA BIND 101 112 BASIC DOMAIN.
CC DOMAIN 113 153 HELIX-LOOP-HELIX MOTIF.
CC DOMAIN 67 75 POLY-GLU.
CC DOMAIN 86 89 POLY-LYS.
CC SEQUENCE 355 AA; 39763 MW; F4344DFD360226B2 CRC64;

Query Match 72.7%; Score 80; DB 1; Length 355;
Best Local Similarity 80.0%; Pred. No. 1.7e-05;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ANARERRMHGHNHAFDQLR 21
| | | | | | | | | | | | | | | | | |
Db 105 ANARERNMHGNAALDNLRL 124

RESULT 12
NDF1 HUMAN
ID NDF1 HUMAN STANDARD; PRT; 356 AA.
AC Q13562; Q00343; Q13340; Q96TH0; Q99455;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neurogenic differentiation factor 1 (Neurod1) (Neurod).
GN NEUROD1 OR NEUROD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96374836; PubMed=8786144;
RA Tamimi R., Steingrimsson E., Copeland N.G., Dyer-Montgomery K.,
RA Lee J.E., Hernandez R., Jenkins N.A., Tapscott S.J.;
RT "The NEUROD gene maps to human chromosome 2q32 and mouse chromosome
RT 2.";
RL Genomics 34:418-421 (1996).
RN [2]

```


RC STRAIN=White leghorn; TISSUE=Retina;
 RA MEDLINE=97454246; PubMed=9310321;
 RX Roztocil T., Matter-Sadzinski L., Alliod C., Ballivet M., Matter J.M.;
 RT "NeuroM, a neural helix-loop-helix transcription factor, defines a
 RL new transition stage in neurogenesis.";
 RN Development 124:3263-3272(1997).
 RP SEQUENCE FROM N.A.
 RC STRAIN=White leghorn; TISSUE=Brain;
 RX MEDLINE=98410737; PubMed=9740021;
 RA Yan R.-T., Wang S.-Z.;
 RT "NeuroD induces photoreceptor cell overproduction in vivo and de novo
 RL generation in vitro.";
 RC J. Neurobiol. 36:485-496(1998).
 CC -!- FUNCTION: Acts as a differentiation factor during neurogenesis.
 CC Induces photoreceptor cell overproduction in vivo and de novo
 CC production in vitro. May play a role in photoreceptor cell
 CC production. Transcriptional activator. Binds DNA on e-box
 CC consensus sequence 5'-CANNTG-3'.
 CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
 CC bHLH protein.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -!- DEVELOPMENTAL STAGE: In the spinal cord it is found in neurons
 CC that are migrating or have reached their final position. Expressed
 CC in cells located at the outer portion of the developing retinal
 CC neuroepithelium, the location where prospective photoreceptors
 CC reside.
 CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Y09596; CAA70784.1; -;
 CC EMBL; AF060885; AAC79425.1; -;
 CC InterPro: IPR001092; HLH_bas.
 CC Pfam; PF00010; HLH; 1.
 CC SMART; SM00353; HLH; 1.
 CC PROSITE; PS50888; HLH; 1.
 CC DNA-binding; Nuclear protein; Transcription regulation; Activator;
 CC Neurogenesis; Developmental protein; Differentiation.
 CC FT DOMAIN 58 77 GLU-RICH (ACIDIC).
 CC FT DOMAIN 90 96 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 CC FT DNA BIND 105 116 BASIC DOMAIN.
 CC FT DOMAIN 117 157 HELIX-LOOP-HELIX MOTIF.
 CC FT DOMAIN 69 78 POLY-GLU.
 CC FT DOMAIN 90 93 POLY-LYS.
 CC FT CONFLICT 282 282 N -> Y (IN REF. 2).
 CC FT CONFLICT 297 297 K -> N (IN REF. 2).
 CC FT CONFLICT 331 331 G -> A (IN REF. 2).
 CC SQ SEQUENCE 357 AA; 38809 MW; 770649CF9DDC54F6 CRC64;
 Query Match 72.7%; Score 80; DB 1; Length 357;
 Best Local Similarity 80.0%; Pred. No. 1.7e-05;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 2 ANARERRRMHGLNHAFDOLR 21
 NDFF1_MOUSE
 Db 109 ANARERRRMHGLNAAALDNL 128
 RESULT 14
 ID NDFF1_MOUSE
 AC Q60857; Q60897; PRT; 357 AA.
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Neurogenic differentiation factor 1 (NeuroD1).
 DE Rattus norvegicus (Rat).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 NC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MFL and 129/SV;
 RX MEDLINE=95273957; PubMed=7754368;
 RA Lee J.E., Hollenberg S.M., Snider L., Turner D.L., Lipnick N.,
 RA Weintraub H.;
 RT "Conversion of Xenopus ectoderm into neurons by NeuroD, a basic
 RL helix-loop-helix protein.";
 CC Science 268:836-844(1995).
 CC -!- FUNCTION: Acts as a differentiation factor during neurogenesis.
 CC Transcriptional activator. Binds to the insulin gene E-Box.
 CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
 CC bHLH protein. Heterodimer with E47.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -!- TISSUE SPECIFICITY: Expressed in differentiating neurons of
 CC both the central and peripheral nervous systems.
 CC -!- DEVELOPMENTAL STAGE: Expressed during embryonic development.
 CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U28068; AAC52203.1; -;
 CC EMBL; U28888; AAC52204.1; -;
 CC PIR; I49338; I49338.
 CC MGD; MGI:133708; Neurod1.
 CC InterPro: IPR001092; HLH_bas.
 CC Pfam; PF00010; HLH; 1.
 CC SMART; SM00353; HLH; 1.
 CC PROSITE; PS50888; HLH; 1.
 CC DNA-binding; Nuclear protein; Transcription regulation; Activator;
 CC Neurogenesis; Developmental protein; Differentiation.
 CC FT DOMAIN 58 77 GLU-RICH (ACIDIC).
 CC FT DOMAIN 87 93 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 CC FT DNA BIND 102 113 BASIC DOMAIN.
 CC FT DOMAIN 114 154 HELIX-LOOP-HELIX MOTIF.
 CC FT DOMAIN 58 64 POLY-GLU.
 CC FT DOMAIN 67 77 POLY-GLU.
 CC FT DOMAIN 87 90 POLY-LYS.
 CC SQ SEQUENCE 357 AA; 39998 MW; B6626E1315E31027 CRC64;
 Query Match 72.7%; Score 80; DB 1; Length 357;
 Best Local Similarity 80.0%; Pred. No. 1.7e-05;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 2 ANARERRRMHGLNHAFDOLR 21
 NDFF1_RAT
 Db 106 ANARERRRMHGLNAAALDNL 125
 RESULT 15
 ID NDFF1_RAT
 AC Q64289;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Neurogenic differentiation factor 1 (NeuroD1) (Basic helix-loop-helix
 DE factor 1) (BHF-1).
 GN NEUROD1 OR NEUROD.
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Cerebellum;
RC MEDLINE=96220182; PubMed=8660336;
RA Kawakami H., Maruyama H., Yasunami M., Ohkubo H., Hara H., Saïda T.,
RA Nakanishi S., Nakamura S.;
RT "Cloning and expression of a rat brain basic helix-loop-helix
RT factor.";
RL Biochem. Biophys. Res. Commun. 221:199-204(1996).
RN [2]
RP SEQUENCE OF 88-200 FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Retina;
RA Ahmad I., Acharay H.R.; to the EMBL/GenBank/DBJ databases.
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Acts as a differentiation factor during neurogenesis.
CC Transcriptional activator. Binds to the insulin gene E-box.
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein. Heterodimer with E47.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D82075; BAA11536.1; -;
DR EMBL; D82074; BAA11535.1; -;
DR EMBL; U80603; AAB38744.1; -;
DR F01; JC4703; JC4703.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS50888; HLH; 1.
KW DNA-binding; Nuclear protein; Transcription regulation; Activator;
KW Neurogenesis; Developmental protein; Differentiation.
FT DOMAIN 58 77 GLU-RICH (ACIDIC).
FT DOMAIN 87 93 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DNA BIND 102 113 BASIC DOMAIN.
FT DOMAIN 114 154 HELIX-LOOP-HELIX MOTIF.
FT DOMAIN 67 76 POLY-GLU.
FT DOMAIN 87 90 POLY-LYS.
SQ SEQUENCE 357 AA; 40000 MW; F773637E64D3E99E CRC64;

Query Match 72.7%; Score 80; DB 1; Length 357;
Best Local Similarity 80.0%; Pred. No. 1.7e-05;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ANARERRMHGLNHAFTQLR 21
||| ||| ||| ||| |||
Db 106 ANARERNMHGLNAALDNL 125

Search completed: September 21, 2004, 21:21:40
Job time : 7.624 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 21:04:52 ; Search time 8.12 Seconds
(without alignments)
815.995 Million cell updates/sec

Title: US-09-980-381A-70
Perfect score: 110
Sequence: 1 AANARERRRMEGLNHAFDQLR 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	64	13	Q8JHA1
2	110	100.0	161	13	Q98939
3	110	100.0	177	13	Q80005
4	107	97.3	266	13	Q42245
5	98	89.1	46	13	Q98904
6	91	82.7	45	5	Q94791
7	88	80.0	134	13	Q8JHA0
8	88	80.0	149	11	Q922E5
9	88	80.0	151	13	Q57598
10	88	80.0	152	4	Q8N100
11	87	79.1	138	13	Q13126
12	87	79.1	138	13	Q13125
13	85	77.3	134	13	Q9DGA9
14	85	77.3	134	13	Q8AW52
15	84	76.4	45	5	Q94790
16	83	75.5	202	5	Q8T360

17	83	75.5	325	13	Q9W6C6
18	82	74.5	189	5	Q9XZC7
19	82	74.5	189	5	Q9V7M2
20	81	73.6	337	4	Q8IYR9
21	81	73.6	337	4	Q7Z4W9
22	80	72.7	216	13	Q9W6B8
23	80	72.7	255	4	Q8IWS6
24	80	72.7	347	13	Q9DE43
25	80	72.7	347	13	Q8JH35
26	80	72.7	356	4	Q9UEC8
27	80	72.7	357	11	Q8CEI7
28	79	71.8	166	4	Q96RJ6
29	79	71.8	168	11	Q923Z4
30	79	71.8	336	13	Q9DDQ8
31	78	70.9	195	5	Q9VGJ5
32	77	70.0	151	13	Q8AX95
33	77	70.0	232	5	Q9VJCI
34	77	70.0	238	13	Q7T301
35	77	70.0	350	13	Q42202
36	77	70.0	355	11	Q8C6A8
37	77	70.0	355	11	Q9JL05
38	77	70.0	381	4	Q8NFK8
39	76	69.1	316	13	Q9W6C7
40	73	66.4	220	13	Q8JIS0
41	73	66.4	223	11	Q8BGW3
42	73	66.4	225	4	Q8NDY6
43	73	66.4	255	5	Q9NB27
44	73	66.4	381	4	Q9UQC6
45	73	65.4	382	4	Q8TBI7

ALIGNMENTS

RESULT 1

Q8JHA1 PRELIMINARY; PRT; 64 AA.

AC Q8JHA1;

DT 01-OCT-2002 (Tremblrel. 22, Created)

DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)

DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)

DE BHLH transcription factor ath1 (Fragment).

OS Serinus canaria (Canary).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Passeriformes; Passeroidea;

OC Fringillidae; Carduelinae; Serinus.

OX NCBI_TaxID=91135;

RN [1]

RP SEQUENCE FROM N.A.

RA Huverstuhl J., Brors D., Bodmer D., Mullen L., Gleich O., Strutz J.,

RA Ryan A.F.;

RT "Expression of BHLH and class IV POU-domain transcription factors in

RT the chicken and canary inner ear";

RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF

CC TRANSCRIPTION FACTORS.

DR EMBL; AF526421; AAM89247.1; -.

DR InterPro; IPR001092; HLH_basic.

DR Pfam; PF00010; HLH; 1.

DR SMART; SMO0353; HLH; 1.

DR PROSITE; PS50888; HLH_2; 1.

FT NON_TER 1

FT NON_TER 64

SQ SEQUENCE 64 AA; 7439 MW; BB691E056670CAB4 CRC64;

Query Match 100.0%; Score 110; DB 13; Length 64;

Best Local Similarity 100.0%; Pred. No. 4.4e-10;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AANARERRRMEGLNHAFDQLR 21

Db 11 AANARERRRMEGLNHAFDQLR 31

```

NON_IER
SEQUENCE
177 AA; 19170 MW; FD44269C5994F5E9 CRC64;

```


DR InterPro; IPR001092; HLH_basic.
 DR Pfam; PF00010; HLH; 1.
 DR SMART; SM00353; HLH; 1.
 DR PROSITE; PS00888; HLH 2; 1.
 FT NON_TER 1
 FT NON_TER 46
 SQ SEQUENCE 46 AA; 5422 MW; 4F88FFD41319483C CRC64;

Query Match 89.1%; Score 98; DB 13; Length 46;
 Best Local Similarity 90.5%; Pred. No. 2.6e-08;
 Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AANARERRMEGLNHAFDQLR 21
 |||||
 Db 1 AANARERRMEGLNKAFLDLR 21

RESULT 6
 Q94791 ID Q94791 PRELIMINARY; PRT; 45 AA.

AC Q94791;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE TATH2 (Fragment).

GN TATH2.
 OS Tribolium castaneum (Red flour beetle).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
 OC Tenebrionidae; Tribolium.

OX NCBI_TaxID=7070;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97026280; PubMed=9872459;
 RA Ben-Arie N., McCall A.E., Berkman S., Eichele G., Bellen H.J.,
 Zoghbi H.F.;

RT Evolutionary conservation of sequence and expression of the bHLH
 protein Atonal suggests a conserved role in neurogenesis.";
 RL Hum. Mol. Genet. 5:1207-1216(1996).

CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
 TRANSCRIPTION FACTORS.

CC EMBL; U61152; AAB41307.1; -.
 DR InterPro; IPR001092; HLH_basic.
 DR Pfam; PF00010; HLH; 1.
 DR SMART; SM00353; HLH; 1.
 DR PROSITE; PS00888; HLH 2; 1.
 FT NON_TER 1
 FT NON_TER 45
 SQ SEQUENCE 45 AA; 5263 MW; 730A6FE821B82E23 CRC64;

Query Match 82.7%; Score 91; DB 5; Length 45;
 Best Local Similarity 85.7%; Pred. No. 3.3e-07;
 Matches 18; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AANARERRMEGLNHAFDQLR 21
 |||||
 Db 1 AANARERRMEGLNEAFDLR 21

RESULT 7
 Q8UHAO ID Q8UHAO PRELIMINARY; PRT; 134 AA.

AC Q8UHAO;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE BHLH transcription factor at5 (Fragment).

OS Serinus canaria (Canary).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Passeriformes; Passeroidea;
 OC Fringillidae; Carduelinae; Serinus.
 OX NCBI_TaxID=9135;
 RN [1]

RP SEQUENCE FROM N.A.
 RA Huverstuhl J., Brors D., Bodmer D., Mullen L., Gleich O., Strutz J.,
 RA Ryan A.F.;
 RT "Expression of bHLH and class IV POU-domain transcription factors in
 the chicken and canary inner ear.";
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
 TRANSCRIPTION FACTORS.

CC EMBL; AF526422; AAM89248.1; -.
 DR InterPro; IPR001092; HLH_basic.
 DR Pfam; PF00010; HLH; 1.
 DR SMART; SM00353; HLH; 1.
 DR PROSITE; PS00888; HLH 2; 1.
 FT NON_TER 134
 FT NON_TER 134
 SQ SEQUENCE 134 AA; 15448 MW; 0016C7C6807F0D6C CRC64;

Query Match 80.0%; Score 88; DB 13; Length 134;
 Best Local Similarity 85.7%; Pred. No. 3.1e-06;
 Matches 18; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AANARERRMEGLNHAFDQLR 21
 |||||
 Db 43 AANARERRMQGLNTAFDLR 63

RESULT 8

Q9Z2E5 ID Q9Z2E5 PRELIMINARY; PRT; 149 AA.

AC Q9Z2E5;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Transcription factor MATH5 (ATOH7).

GN ATOH7.
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=129SV/J; TISSUE=Retina;
 RX MEDLINE=99026071; PubMed=9806930;
 RA Brown N.L., Kanekar S., Vetter M.L., Tucker P.K., Gemza D.L.,
 RA Glaser T.;
 RT "Math5 encodes a murine basic helix-loop-helix transcription factor
 expressed during early stages of retinal neurogenesis.";
 RL Development 125:4821-4833(1998).
 RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=129S4;
 RX MEDLINE=1886645; PubMed=11889557;
 RA Brown N.L., Dagenais S.L., Chen C.M., Glaser T.;
 RT "Molecular characterization and mapping of ATOH7, a human atonal
 homolog with a predicted role in retinal ganglion cell development.";
 RL Mamm. Genome 13:95-101(2002).
 DR EMBL; AF071223; AAC68868.1; -.
 DR EMBL; AF418923; AAL11912.1; -.
 DR MGD; MGI:135553; Atoh7.
 DR GO; GO:0009649; P:entrainment of circadian clock; IMP.
 DR InterPro; IPR001092; HLH_basic.
 DR Pfam; PF00010; HLH; 1.
 DR PROSITE; PS00888; HLH 2; 1.

SQ SEQUENCE 149 AA; 16569 MW; DF78A25190B5FD3A CRC64;

Query Match 80.0%; Score 88; DB 11; Length 149;
 Best Local Similarity 85.7%; Pred. No. 3.4e-06;
 Matches 18; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AANARERRMEGLNHAFDQLR 21
 |||||
 Db 45 AANARERRMQGLNTAFDLR 65

```

RESULT 9
O57598 PRELIMINARY; PRT; 151 AA.
AC O57598;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Atonal transcription factor homologue.
GN ATH5.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Neuroretina; PubMed=11124117;
RX MEDLINE=21064448; PubMed=11124117;
RA Matter-Sadzinski L., Matter J.M., Ong M.T., Hernandez J., Ballivet M.;
RT "Specification of neurotransmitter receptor identity in developing
RT retina: the chick ATH5 promoter integrates the positive and negative
RT effects of several bHLH proteins.";
RL Development 128:217-231 (2001).
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR EMBL; AJ001178; CAA04572.1; -.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00888; HLH 2; 1.
FT CHAIN 2 151 ATONAL TRANSCRIPTION FACTOR HOMOLOGUE.
SQ SEQUENCE 151 AA; 17540 MW; 1D4F2F9D225A5319 CRC64;

Query Match 80.0%; Score 88; DB 13; Length 151;
Best Local Similarity 85.7%; Pred. No. 3.4e-06;
Matches 18; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AANARERRRMHGLNHAFDQLR 21
| | | | | | | | | | | | | | | | | | | |
DB 43 AANARERRRMQGLNTAFDRLR 63
| | | | | | | | | | | | | | | | | | | |

RESULT 10
O8N100 PRELIMINARY; PRT; 152 AA.
AC O8N100;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Atonal homolog 7 (ATOH7).
GN ATOH7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Eye;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21886645; PubMed=11889557;
RA Brown N.L., Degenais S.L., Chen C.M., Glaser T.;
RT "Molecular characterization and mapping of ATOH7, a human atonal
RT homolog with a predicted role in retinal ganglion cell development.";
RL Mamm. Genome 13:95-101 (2002).
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR EMBL; BC032621; AAB32621.1; -.
DR EMBL; AF418922; AAL11911.1; -.
DR Genew; HGNC:113907; ATOH7.
DR InterPro; IPR001092; HLH_basic.

```

```

DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00888; HLH 2; 1.
SQ SEQUENCE 152 AA; 16871 MW; 9E93E9E0E1697C5 CRC64;

Query Match 80.0%; Score 88; DB 4; Length 152;
Best Local Similarity 85.7%; Pred. No. 3.5e-06;
Matches 18; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AANARERRRMHGLNHAFDQLR 21
| | | | | | | | | | | | | | | | | | | |
DB 44 AANARERRRMQGLNTAFDRLR 64
| | | | | | | | | | | | | | | | | | | |

RESULT 11
O13126 PRELIMINARY; PRT; 138 AA.
AC O13126;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Atonal homolog 5b.
GN XATH5B.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Head;
RA Kanekar S., Perron M., Harris W.A., Jan L.Y., Jan Y.N., Vetter M.L.;
RT "Xath5 participates in a network of bHLH genes in the developing
RT Xenopus retina.";
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR EMBL; U93171; AAB58669.1; -.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00888; HLH 2; 1.
SQ SEQUENCE 138 AA; 16028 MW; A72A5DD9BE118EDE CRC64;

Query Match 79.1%; Score 87; DB 13; Length 138;
Best Local Similarity 85.7%; Pred. No. 4.5e-06;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AANARERRRMHGLNHAFDQLR 21
| | | | | | | | | | | | | | | | | | | |
DB 37 AANARERRRMQGLNTAFDLSLR 57
| | | | | | | | | | | | | | | | | | | |

RESULT 12
O13125 PRELIMINARY; PRT; 138 AA.
AC O13125;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Atonal homolog 5a.
GN XATH5A.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Head;
RA Kanekar S., Perron M., Harris W.A., Jan L.Y., Jan Y.N., Vetter M.L.;
RT "Xath5 participates in a network of bHLH genes in the developing
RT Xenopus retina.";

```

```

RN RP SEQUENCE FROM N.A.
RA Lloyd D.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL627094; CAD52125.1; -.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS50888; HLH_2; 1.
SQ SEQUENCE 134 AA; 15621 MW; 3207DBD0E5972496 CRC64;

Query Match 77.3%; Score 85; DB 13; Length 134;
Best Local Similarity 81.0%; Pred. No. 9.2e-06;
Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps

Qy 1 AANARERRRMHGLNHAFDQLR 21
   |||||:|||||:|||||
Db 32 AANAREKRMQGLNTAFDLRL 52

RESULT 15
Q94790 PRELIMINARY; PRT; 45 AA.
ID Q94790
AC Q94790;
DT 01-FEB-1997 (TReMBLrel. 02, Created)
DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE TATH1 (Fragment).
GN TATH1
OS Tribolium castaneum (Red flour beetle).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
OC Tenebrionidae; Tribolium.
OX NCBI_TaxID=7070;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97026280; PubMed=8872459;
RA Zoghbi H.Y.;
RA "Evolutionary conservation of sequence and expression of the BHLH
RT protein Atonal suggests a conserved role in neurogenesis.";
RL Hum. Mol. Genet. 5:1207-1216(1996).
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSSCRIPTION FACTORS.
DR EMBL; U61151; AAB41306.1; -.
DR HSP; P25912; 1HLO.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS50888; HLH_2; 1.
FT NON TER 1
FT NON TER 45
SQ SEQUENCE 45 AA; 5256 MW; 633ACFB9E1FE5F0A CRC64;

Query Match 76.4%; Score 84; DB 5; Length 45;
Best Local Similarity 81.0%; Pred. No. 4.4e-06;
Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps

Qy 1 AANARERRRMHGLNHAFDQLR 21
   |||||:|||||:|||||
Db 1 AANAREKRMNSLINDAFDLRL 21

Search completed: September 21, 2004, 21:24:12
Job time : 9.12 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2004, 12:22:55 ; Search time 0.001 Seconds
(without alignments)
10.710 Million cell updates/sec

Title: us-10-084-555a-115
Perfect score: 21
Sequence: 1 ttgtgtggggagtatttgagt 21

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 0.5

Searched: 1 seqs, 255 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 2000 summaries

Database : us-10-084-555a-8:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	6.4	30.5	255	1	us-10-084-555a-8
2	5	23.8	255	1	us-10-084-555a-8

ALIGNMENTS

RESULT 1
us-10-084-555a-8/c

Query Match 30.5%; Score 6.4; DB 1; Length 255;
Best Local Similarity 62.5%; Pred. No. 0;
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 TGTGGGGAGTTATTGA 19
||| ||| ||| |||
Db 200 TGCAGGAGAGAGAGGA 185

RESULT 2
us-10-084-555a-8

Query Match 23.8%; Score 5; DB 1; Length 255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 TGAGT 21
||| ||| |||
Db 22 TGAGT 26

Search completed: September 22, 2004, 12:22:55
Job time : 0.001 secs

Search completed: September 22, 2004, 12:23:44
Job time : 0.001 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2004, 12:24:40 ; Search time 0.001 Seconds
(without alignments)
9.180 Million cell updates/sec

Title: us-10-084-555a-117
Perfect score: 18
Sequence: 1 tgtggggagttatcgagc 18

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 0.5

Searched: 1 seqs, 255 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 2000 summaries

Database : us-10-084-555a-8:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	7.6	42.2	255	1	us-10-084-555a-8
2	5	27.8	255	1	us-10-084-555a-8

ALIGNMENTS

RESULT 1
us-10-084-555a-8/c

Query Match 42.2%; Score 7.6; DB 1; Length 255;
Best Local Similarity 71.4%; Pred. No. 0;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 GGGAGTTATCGAGC 18
||| |||||
Db 113 GGGGATCGTCGAGC 100

RESULT 2
us-10-084-555a-8

Query Match 27.8%; Score 5; DB 1; Length 255;
Best Local Similarity 61.5%; Pred. No. 0;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 6 GGAGTTATCGAGC 18
||| |||||
Db 79 GGAGCATTCTGC 91

Search completed: September 22, 2004, 12:24:41
Job time : 1 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2004, 12:25:32 ; Search time 0.001 Seconds
(without alignments)
9.690 Million cell updates/sec

Title: us-10-084-555a-118
Perfect score: 19
Sequence: 1 gccttcgcgaaaaaaatcg 19

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 0.5

Searched: 1 seqs, 255 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 2000 summaries

Database : us-10-084-555a-8:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	7.6	40.0	255	1	us-10-084-555a-8
2	5.8	30.5	255	1	us-10-084-555a-8

ALIGNMENTS

RESULT 1
us-10-084-555a-8/c

Query Match 40.0%; Score 7.6; DB 1; Length 255;
Best Local Similarity 71.4%; Pred. No. 0;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GCCTTCGCGAAAA 14
||| ||| ||| |||
Db 168 GCCTGGGGCGCAGAA 155

RESULT 2
us-10-084-555a-8

Query Match 30.5%; Score 5.8; DB 1; Length 255;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCTTCGCG 9
||| ||| ||| |||
Db 28 GCCTCCACG 36

Search completed: September 22, 2004, 12:25:32
Job time : 0.001 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 20:34:21 ; Search time 9.912 Seconds
(without alignments)
598.618 Million cell updates/sec

Title: US-09-980-381A-70

Perfect score: 110

Sequence: 1 AANARERRMEGLNHPDQLR 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 29Jan04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	21	4	AAB60379
2	110	100.0	21	6	AAC030950
3	110	100.0	161	4	AAB60376
4	110	100.0	161	6	AAC030946
5	110	100.0	259	6	AAC030949
6	110	100.0	351	4	AAB60353
7	110	100.0	351	4	AAB60369
8	110	100.0	351	6	AAC030939
9	110	100.0	351	6	AAC030924
10	110	100.0	354	4	AAB60375
11	110	100.0	354	4	AAB60349
12	110	100.0	354	6	AAC030945
13	110	100.0	354	6	AAC030920
14	107	97.3	266	4	AAB60366
15	107	97.3	266	6	AAC030936
16	98	89.1	46	4	AAB60377
17	98	89.1	46	6	AAC030947
18	91	82.7	45	4	AAB60373
19	91	82.7	45	6	AAC030943
20	88	80.0	149	4	AAB60358
21	88	80.0	149	6	AAC030928
22	88	80.0	151	4	AAB60364
23	88	80.0	151	6	AAC030934
24	87	79.1	138	4	AAB60372
25	87	79.1	138	4	AAB60371

ALIGNMENTS

RESULT 1

AAB60379
ID AAB60379 standard; peptide; 21 AA.

XX AAB60379;

DT 24-APR-2001 (first entry)

DE Atonal-associated peptide fragment.

XX

Atonal; homologue; orthologue; atonal-associated protein; deafness; hearing impairment; vestibular effect; balance disorder; osteoarthritis; cellular proliferation; cerebellar granule neuron; gene therapy; mechanoreceptive cell growth; auditory; osteopathic; cytostatic; transgenic animal.

XX Unidentified.

XX WO200073764-A2.

XX 07-DEC-2000.

XX 01-JUN-2000; 2000WO-US015410.

XX 01-JUN-1999; 99US-0137060P.

XX 19-JAN-2000; 2000US-0176993P.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX Zoghbi HY, Bellen H, Birmingham N, Hassan B, Ben-Arie N;

XX WPI; 2001-032190/04.

XX Therapeutic use of atonal-associated nucleic acids or amino acids, or any of its homologs or orthologs, for the treatment of e.g. deafness, osteoarthritis and abnormal cell proliferation.

XX Disclosure; Page 33; 142pp; English.

XX The invention relates to the use of atonal-associated nucleic acid or amino acid sequence, or any of its homologs or orthologues as therapeutic agents for the treatment of deafness, partial hearing loss, vestibular effects due to damage or loss of inner hair cells, osteoarthritis and abnormal cell proliferation. The invention also encompasses methods of screening for compounds which affect the expression of an atonal-associated nucleic acid sequence in an animal, and a transgenic animal in which an allele of a native atonal-associated gene is replaced by a heterologous nucleic acid sequence, thus

26 87 79.1 138 6 AAC030942
27 87 79.1 138 6 AAC030941
28 84 76.4 45 4 AAB60374
29 84 76.4 45 6 AAC030944
30 84 76.4 198 4 ABB71919
31 84 76.4 312 4 ABB63330
32 84 76.4 312 4 AAB60378
33 84 76.4 312 6 AAC030948
34 82 74.5 189 4 ABB63483
35 82 74.5 189 4 AAB60355
36 82 74.5 189 6 AAC030925
37 81 73.6 245 4 AAM41883
38 81 73.6 337 3 AAB12503
39 81 73.6 337 4 AAG66722
40 81 73.6 337 4 ABB50169
41 81 73.6 337 4 AAB60352
42 81 73.6 337 4 AAB60368
43 81 73.6 337 6 AAC030938
44 81 73.6 337 6 AAC030923
45 81 73.6 352 2 AAR77505

AAC030942 Frog aton
AAC030941 Frog aton
AAB60374 T. casten
AAC030944 Beetle at
ABB71919 Drosophil
ABB63330 Drosophil
AAB60378 Drosophil
AAC030948 Drosophil
ABB63483 Drosophil
AAB60355 Drosophil
AAC030925 Drosophil
AAM41883 Human pol
AAB12503 Human MAT
AAG66722 Human ato
ABB50169 Human tra
AAB60352 Mouse ato
AAB60368 Mouse ato
AAC030938 Mouse ato
AAC030923 Mouse ato
AAR77505 Frog neur

CC inactivating the atonal-associated allele. The nucleic acids or proteins
 CC may be used in a method of treating an animal for hearing impairment,
 CC joint disease, balance disorders, abnormal cell proliferation, or other
 CC disease related to loss of a functional atonal-associated nucleic acid or
 CC protein. They may particularly be used to treat an animal with a
 CC deficiency in cerebellar granule neurons or their precursors, and may
 CC also be used in promoting mechanoreceptive cell growth and generating
 CC hair cells. The present sequence represents an atonal-associated amino
 CC acid sequence referred to in the invention
 XX
 SQ Sequence 21 AA;

Query Match 100.0%; Score 110; DB 4; Length 21;
 Best Local Similarity 100.0%; Pred. No. 9.1e-12;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRMHGLNHAFDQLR 21
 |||||
 Db 1 AANARERRMHGLNHAFDQLR 21

RESULT 2
 AAO30950
 ID AAO30950 standard; peptide; 21 AA.

AC AAO30950;

XX 22-SEP-2003 (first entry)

DE Transcription factor peptide used in the invention.

XX Atonal-associated protein; gastrointestinal; damaged intestinal tissue;
 KW cancer; inflammatory bowel disease; irritable bowel syndrome; infection;
 KW necrotising enterocolitis; hearing impairment; abnormal proliferation;
 KW joint disease; gene therapy; imbalance disorder; transcription factor.

OS Unidentified.

XX WO2003047532-A2.

PN 12-JUN-2003.

XX 03-DEC-2002; 2002WO-US041458.

PF 05-DEC-2001; 2001US-00004717.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX Zoghbi HY, Qi Y;

XX WPI; 2003-505253/47.

XX New compositions comprising at least one stem cell which is up- or down-
 PT regulated for expression of an atonal-associated sequence, and at least
 PT one regulatory factor, useful for treating a gastrointestinal condition,
 PT e.g. cancer.

PS Disclosure; Page 27; 157pp; English.

XX The invention relates to pharmaceutical compositions and methods for the
 CC therapeutic use of an atonal-associated sequence. The composition
 CC comprises of a stem cell which is up- or down-regulated for expression of
 CC an atonal-associated sequence and a regulatory factor. Atonal-associated
 CC amino acid or nucleic acid sequence is used to treat a gastrointestinal
 CC condition including cancer, damaged intestinal tissue, inflammatory bowel
 CC disease, irritable bowel syndrome, infection or necrotising enterocolitis.
 CC They are also useful for promoting mechanoreceptive cell growth, for
 CC generating hair cells, for treating hearing impairment or imbalance
 CC disorder, joint disease, abnormal proliferation (e.g. cancer), or a
 CC disease that is a result or loss of functional atonal-associated
 CC sequences. Atonal-associated sequences are also used in gene therapy. The
 CC present sequence is a transcription factor peptide used to illustrate the
 CC method of the invention

XX
 SQ

Sequence 21 AA;

Query Match 100.0%; Score 110; DB 6; Length 21;
 Best Local Similarity 100.0%; Pred. No. 9.1e-12;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRMHGLNHAFDQLR 21
 |||||
 Db 1 AANARERRMHGLNHAFDQLR 21

RESULT 3
 AAB60376
 ID AAB60376 standard; protein; 161 AA.

AC AAB60376;

XX 24-APR-2001 (first entry)

XX Chicken atonal homologue 1 (Cath1) protein, SEQ ID NO:60.

XX Atonal; homologue; orthologue; atonal-associated protein; deafness;
 KW hearing impairment; vestibular effect; balance disorder; osteoarthritis;
 KW cellular proliferation; cerebellar granule neuron; gene therapy;
 KW mechanoreceptive cell growth; auditory; osteopathic; cytostatic;
 KW transgenic animal.

XX Gallus gallus.

OS WO200073764-A2.

XX 07-DEC-2000.

XX 01-JUN-2000; 2000WO-US015410.

XX 01-JUN-1999; 99US-0137060P.

XX 19-JAN-2000; 2000US-0176993P.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX Zoghbi HY, Bellen H, Birmingham N, Hassan B, Ben-Arie N;

XX WPI; 2001-032190/04.

XX N-PSDB; AAF27284.

XX Therapeutic use of atonal-associated nucleic acids or amino acids, or any
 PT of its homologs or orthologs, for the treatment of e.g. deafness,
 PT osteoarthritis and abnormal cell proliferation.

XX Disclosure; Page; 142pp; English.

XX The invention relates to the use of atonal-associated nucleic acid or
 CC amino acid sequence, or any of its homologues or orthologues as
 CC therapeutic agents for the treatment of deafness, partial hearing loss,
 CC vestibular effects due to damage or loss of inner hair cells,
 CC osteoarthritis and abnormal cell proliferation. The invention also
 CC encompasses methods of screening for compounds which affect the
 CC expression of an atonal-associated nucleic acid sequence in an animal,
 CC and a transgenic animal in which an allele of a native atonal-associated
 CC gene is replaced by a heterologous nucleic acid sequence, thus
 CC inactivating the atonal-associated allele. The nucleic acids or proteins
 CC may be used in a method of treating an animal for hearing impairment,
 CC joint disease, balance disorders, abnormal cell proliferation, or other
 CC disease related to loss of a functional atonal-associated nucleic acid or
 CC protein. They may particularly be used to treat an animal with a
 CC deficiency in cerebellar granule neurons or their precursors, and may
 CC also be used in promoting mechanoreceptive cell growth and generating
 CC hair cells. The present sequence represents an atonal-associated amino
 CC acid sequence referred to in the invention. Note: The present sequence is
 CC not shown in the specification, but was obtained from GenBank

XX Sequence 161 AA;

Query Match 100.0%; Score 110; DB 4; Length 161;
 Best Local Similarity 100.0%; Pred. No. 8.6e-11;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRMHGLNHAFLDLR 21
 |||||
 Db 89 AANARERRMHGLNHAFLDLR 109

RESULT 4
 AAO30946
 ID AAO30946 standard; protein; 161 AA.
 AC AAO30946;
 XX
 DT 22-SEP-2003 (first entry)
 XX
 DE Chicken atonal homologue 1 (Cath1) protein #4.
 XX

KW Atonal-associated protein; gastrointestinal; damaged intestinal tissue;
 KW cancer; inflammatory bowel disease; irritable bowel syndrome; infection;
 KW necrotising enterocolitis; hearing impairment; abnormal proliferation;
 KW joint disease; gene therapy; imbalance disorder; atonal homologue 1;
 KW Cath1; chicken.
 XX

OS Gallus sp.
 XX

PN WO2003047532-A2.
 XX

PD 12-JUN-2003.
 XX

PF 03-DEC-2002; 2002WO-US041459.
 XX

PR 05-DEC-2001; 2001US-00004717.
 XX

PA (BAYU) BAYLOR COLLEGE MEDICINE.
 XX

PI Zoghbi HY, Qi Y;
 XX

DR WPI; 2003-505253/47.
 XX

DR N-PSDB; AAL62231.
 XX

XX New compositions comprising at least one stem cell which is up- or down-
 regulated for expression of an atonal-associated sequence, and at least
 one regulatory factor, useful for treating a gastrointestinal condition,
 e.g. cancer.
 XX

PS Disclosure; Page 152-153; 157pp; English.
 XX

XX The invention relates to pharmaceutical compositions and methods for the
 CC therapeutic use of an atonal-associated sequence. The composition
 CC comprises a stem cell which is up- or down-regulated for expression of
 CC an atonal-associated sequence and a regulatory factor. Atonal-associated
 CC amino acid or nucleic acid sequence is used to treat a gastrointestinal
 CC condition including cancer, damaged intestinal tissue, inflammatory bowel
 CC disease, irritable bowel syndrome, infection or necrotising enterocolitis.
 CC They are also useful for promoting mechanoreceptive cell growth, for
 CC generating hair cells, for treating hearing impairment or imbalance
 CC disorder, joint disease, abnormal proliferation (e.g. cancer), or a
 CC disease that is a result or loss of functional atonal-associated
 CC sequences. Atonal-associated sequences are also used in gene therapy. The
 CC present sequence is chicken atonal homologue 1 (Cath1) protein. This
 CC sequence is used to illustrate the method of the invention
 XX

SQ Sequence 161 AA;
 XX

Query Match 100.0%; Score 110; DB 6; Length 161;
 Best Local Similarity 100.0%; Pred. No. 8.6e-11;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRMHGLNHAFLDLR 21
 |||||

Db 89 AANARERRMHGLNHAFLDLR 109

RESULT 5

AAO30949

ID AAO30949 standard; protein; 259 AA.

AC AAO30949;

XX

DT 22-SEP-2003 (first entry)

XX

DE Frog atonal homologue 1 (Xath1) protein #4.

XX

KW Atonal-associated protein; gastrointestinal; damaged intestinal tissue;
 KW cancer; inflammatory bowel disease; irritable bowel syndrome; infection;
 KW necrotising enterocolitis; hearing impairment; abnormal proliferation;
 KW joint disease; gene therapy; imbalance disorder; atonal homologue 1;
 KW Xath1; frog.
 KW

OS Xenopus sp.

XX

XX Key Location/Qualifiers

FT Misc-difference 61..62 /note= "Encoded by GCCGGCGGCAC"

FT Misc-difference 81..83 /note= "Encoded by GAGGAT"

FT Misc-difference 197..199 /note= "Encoded by AGCATG"

FT Misc-difference 203 /note= "Encoded by GAG"

FT Misc-difference 204 /note= "Encoded by ACT"

FT Misc-difference 205 /note= "Encoded by TCC"

FT Misc-difference 207 /note= "Encoded by CGA"

FT Misc-difference 212..214 /note= "Encoded by AGCCAA"

FT Misc-difference 220..224 /note= "Encoded by GCTCCGAGAG"

FT Misc-difference 228..230 /note= "Encoded by TCGGCCCATCTTCATCTTCGGGAGAGACA"

FT TCGGCCCA

FT Misc-difference 237 /note= "Encoded by AGA"

FT Misc-difference 238 /note= "Encoded by AGA"

FT Misc-difference 239 /note= "Encoded by CAG"

FT Misc-difference 240 /note= "Encoded by CAA"

FT Misc-difference 241 /note= "Encoded by GAC"

FT Misc-difference 242 /note= "Encoded by ATC"

FT Misc-difference 243 /note= "Encoded by GCA"

XX WO2003047532-A2.

PD 12-JUN-2003.

XX 03-DEC-2002; 2002WO-US041459.

XX 05-DEC-2001; 2001US-00004717.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX Zoghbi HY, Qi Y;

XX WPI; 2003-505253/47.

XX N-PSDB; AAL62234.

PT New compositions comprising at least one stem cell which is up- or down-regulated for expression of an atonal-associated sequence, and at least one regulatory factor, useful for treating a gastrointestinal condition, e.g. Cancer.

PS Disclosure; Page 155-156; 157pp; English.

XX The invention relates to pharmaceutical compositions and methods for the therapeutic use of an atonal-associated sequence. The composition comprises of a stem cell which is up- or down-regulated for expression of an atonal-associated sequence and a regulatory factor. Atonal-associated amino acid or nucleic acid sequence is used to treat a gastrointestinal condition including cancer, damaged intestinal tissue, inflammatory bowel disease, irritable bowel syndrome, infection or necrotizing enterocolitis. They are also useful for promoting mechanoreceptive cell growth, for generating hair cells, for treating hearing impairment or imbalance disorder, joint disease, abnormal proliferation (e.g. cancer), or a disease that is a result of loss of functional atonal-associated sequences. Atonal-associated sequences are also used in gene therapy. The present sequence is frog atonal homologue 1 (Xathl) protein. This sequence is used to illustrate the method of the invention

SQ Sequence 259 AA;

Query Match 100.0%; Score 110; DB 6; Length 259;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRRMHGLNHAFDQLR 21
|||||
Db 115 AANARERRRMHGLNHAFDQLR 135

RESULT 6
AAB60353
ID AAB60353 standard; protein; 351 AA.

XX

AC AAB60353;

XX 24-APR-2001 (first entry)

DE Mouse atonal homologue 1 (ATOH1, Math1) protein, SEQ ID NO:11.

XX Atonal; homologue; orthologue; atonal-associated protein; deafness;
KW hearing impairment; vestibular effect; balance disorder; osteoarthritis;
KW cellular proliferation; cerebellar granule neuron; gene therapy;
KW mechanoreceptive cell growth; auditory; osteopathic; cytostatic;
KW transgenic animal.

XX Mus musculus.

XX WO200073764-A2.

XX 07-DEC-2000.

XX 01-JUN-2000; 2000WO-US015410.

XX 01-JUN-1999; 99US-0137060P.

XX 19-JAN-2000; 2000US-0176993P.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX Zoghbi HY, Bellen H, Birmingham N, Hassan B, Ben-Arie N;

XX WPI; 2001-032190/04.

XX N-PSDB; AAF27257.

XX Therapeutic use of atonal-associated nucleic acids or amino acids, or any of its homologs or orthologs, for the treatment of e.g. deafness, PT osteoarthritis and abnormal cell proliferation.

PS Disclosure; Page; 142pp; English.

XX

CC The invention relates to the use of atonal-associated nucleic acid or amino acid sequence, or any of its homologues or orthologues as therapeutic agents for the treatment of deafness, partial hearing loss, CC vestibular effects due to damage or loss of inner hair cells, CC osteoarthritis and abnormal cell proliferation. The invention also CC encompasses methods of screening for compounds which affect the CC expression of an atonal-associated nucleic acid sequence in an animal, CC and a transgenic animal in which an allele of a native atonal-associated gene is replaced by a heterologous nucleic acid sequence, thus CC inactivating the atonal-associated allele. The nucleic acids or proteins CC may be used in a method of treating an animal for hearing impairment, CC joint disease, balance disorders, abnormal cell proliferation, or other CC disease related to loss of a functional atonal-associated nucleic acid or CC protein. They may particularly be used to treat an animal with a CC deficiency in cerebellar granule neurons or their precursors, and may CC also be used in promoting mechanoreceptive cell growth and generating CC hair cells. The present sequence represents an atonal-associated amino CC acid sequence referred to in the invention. Note: The present sequence is CC not shown in the specification, but was obtained from GenBank

XX Sequence 351 AA;

Query Match 100.0%; Score 110; DB 4; Length 351;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRRMHGLNHAFDQLR 21

Db 160 AANARERRRMHGLNHAFDQLR 180

RESULT 7

AAB60369

ID AAB60369 standard; protein; 351 AA.

XX

AC AAB60369;

XX 24-APR-2001 (first entry)

DE Mouse atonal homologue 1 (Math-1) protein, SEQ ID NO:46.

XX Atonal; homologue; orthologue; atonal-associated protein; deafness;
KW hearing impairment; vestibular effect; balance disorder; osteoarthritis;
KW cellular proliferation; cerebellar granule neuron; gene therapy;
KW mechanoreceptive cell growth; auditory; osteopathic; cytostatic;
KW transgenic animal.

XX Mus musculus.

XX WO200073764-A2.

XX 07-DEC-2000.

XX 01-JUN-2000; 2000WO-US015410.

XX 01-JUN-1999; 99US-0137060P.

XX 19-JAN-2000; 2000US-0176993P.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX Zoghbi HY, Bellen H, Birmingham N, Hassan B, Ben-Arie N;

XX WPI; 2001-032190/04.

XX N-PSDB; AAF27277.

XX Therapeutic use of atonal-associated nucleic acids or amino acids, or any of its homologs or orthologs, for the treatment of e.g. deafness, PT osteoarthritis and abnormal cell proliferation.

XX Disclosure; Page; 142pp; English.

XX The invention relates to the use of atonal-associated nucleic acid or amino acid sequence, or any of its homologues or orthologues as

therapeutic agents for the treatment of deafness, partial hearing loss, vestibular effects due to damage or loss of inner hair cells, osteoarthritis and abnormal cell proliferation. The invention also encompasses methods of screening for compounds which affect the expression of an atonal-associated nucleic acid sequence in an animal, and a transgenic animal in which an allele of a native atonal-associated gene is replaced by a heterologous nucleic acid sequence, thus inactivating the atonal-associated allele. The nucleic acids or proteins may be used in a method of treating an animal for hearing impairment, joint disease, balance disorders, abnormal cell proliferation, or other disease related to loss of a functional atonal-associated nucleic acid or protein. They may particularly be used to treat an animal with a deficiency in cerebellar granule neurons or their precursors, and may also be used in promoting mechanoreceptive cell growth and generating hair cells. The present sequence represents an atonal-associated amino acid sequence referred to in the invention. Note: The present sequence is not shown in the specification, but was obtained from GenBank

XX Sequence 351 AA;

Query Match 100.0%; Score 110; DB 4; Length 351;
 Best Local Similarity 100.0%; Pred. No. 2e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRRHGLNHAFDQLR 21
 |||||
 Db 160 AANARERRRHGLNHAFDQLR 180

RESULT 8
 AAO30939
 ID AAO30939 standard; protein; 351 AA.

XX AAO30939;

XX 22-SEP-2003 (first entry)

XX Mouse atonal homologue 1 (Math1) protein #13.

XX Atonal-associated protein; gastrointestinal; damaged intestinal tissue;
 KW cancer; inflammatory bowel disease; irritable bowel syndrome; infection;
 KW necrotising enterocolitis; hearing impairment; abnormal proliferation;
 KW joint disease; gene therapy; imbalance disorder; atonal homologue 1;
 KW Math1; mouse.

XX Mus musculus.

XX WO2003047532-A2.

XX 12-JUN-2003.

XX 03-DEC-2002; 2002WO-US041458.

XX 05-DEC-2001; 2001US-00004717.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX Zoghbi HY, Qi Y;

XX WPI; 2003-505253/47.

XX N-PSDB; AAL62224.

XX New compositions comprising at least one stem cell which is up- or down-
 PT regulated for expression of an atonal-associated sequence, and at least
 PT one regulatory factor, useful for treating a gastrointestinal condition,
 PT e.g. cancer.

XX Disclosure; Page 144-145; 157pp; English.

XX The invention relates to pharmaceutical compositions and methods for the
 CC therapeutic use of an atonal-associated sequence. The composition
 CC comprises of a stem cell which is up- or down-regulated for expression of
 CC an atonal-associated sequence and a regulatory factor. Atonal-associated

CC amino acid or nucleic acid sequence is used to treat a gastrointestinal
 CC condition including cancer, damaged intestinal tissue, inflammatory bowel
 CC disease, irritable bowel syndrome, infection or necrotising enterocolitis.
 CC They are also useful for promoting mechanoreceptive cell growth, for
 CC generating hair cells, for treating hearing impairment or imbalance
 CC disorder, joint disease, abnormal proliferation (e.g. cancer), or a
 CC disease that is a result or loss of functional atonal-associated
 CC sequences. Atonal-associated sequences are also used in gene therapy. The
 CC present sequence is mouse atonal homologue 1 (Math1) protein. This
 CC sequence is used to illustrate the method of the invention

XX Sequence 351 AA;

Query Match 100.0%; Score 110; DB 6; Length 351;

Best Local Similarity 100.0%; Pred. No. 2e-10;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRRHGLNHAFDQLR 21
 |||||
 Db 160 AANARERRRHGLNHAFDQLR 180

RESULT 9
 AAO30924

ID AAO30924 standard; protein; 351 AA.

XX AAO30924;

XX 22-SEP-2003 (first entry)

XX Mouse atonal homologue 1 (Math1) protein #4.

XX Atonal-associated protein; gastrointestinal; damaged intestinal tissue;
 KW cancer; inflammatory bowel disease; irritable bowel syndrome; infection;
 KW necrotising enterocolitis; hearing impairment; abnormal proliferation;
 KW joint disease; gene therapy; imbalance disorder; atonal homologue 1;
 KW Math1; mouse.

XX Mus musculus.

XX WO2003047532-A2.

XX 12-JUN-2003.

XX 03-DEC-2002; 2002WO-US041458.

XX 05-DEC-2001; 2001US-00004717.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX Zoghbi HY, Qi Y;

XX WPI; 2003-505253/47.

XX N-PSDB; AAL62204.

XX New compositions comprising at least one stem cell which is up- or down-
 PT regulated for expression of an atonal-associated sequence, and at least
 PT one regulatory factor, useful for treating a gastrointestinal condition,
 PT e.g. cancer.

XX Disclosure; Page 122-123; 157pp; English.

XX The invention relates to pharmaceutical compositions and methods for the
 CC therapeutic use of an atonal-associated sequence. The composition
 CC comprises of a stem cell which is up- or down-regulated for expression of
 CC an atonal-associated sequence and a regulatory factor. Atonal-associated
 CC amino acid or nucleic acid sequence is used to treat a gastrointestinal
 CC condition including cancer, damaged intestinal tissue, inflammatory bowel
 CC disease, irritable bowel syndrome, infection or necrotising enterocolitis.
 CC They are also useful for promoting mechanoreceptive cell growth, for
 CC generating hair cells, for treating hearing impairment or imbalance
 CC disorder, joint disease, abnormal proliferation (e.g. cancer), or a
 CC disease that is a result or loss of functional atonal-associated

The invention relates to the use of atonal-associated nucleic acid or amino acid sequence, or any of its homologues or orthologues as therapeutic agents for the treatment of deafness, partial hearing loss, vestibular effects due to damage or loss of inner hair cells, osteoarthritis and abnormal cell proliferation. The invention also encompasses methods of screening for compounds which affect the expression of an atonal-associated nucleic acid sequence in an animal, and a transgenic animal in which an allele of a native atonal-associated gene is replaced by a heterologous nucleic acid sequence, thus inactivating the atonal-associated allele. The nucleic acids or proteins may be used in a method of treating an animal for hearing impairment, joint disease, balance disorders, abnormal cell proliferation, or other disease related to loss of a functional atonal-associated nucleic acid or protein. They may particularly be used to treat an animal with a deficiency in cerebellar granule neurons or their precursors, and may also be used in promoting mechanoreceptive cell growth and generating hair cells. The present sequence represents an atonal-associated amino acid sequence referred to in the invention. Note: The present sequence is

amino acid sequence, or any of its homologues or orthologues, therapeutic agents for the treatment of deafness, partial hearing loss, vestibular effects due to damage or loss of inner hair cells, osteoarthritis and abnormal cell proliferation. The invention encompasses methods of screening for compounds which affect the expression of an atonal-associated nucleic acid sequence in an animal and a transgenic animal in which an allele of a native atonal gene is replaced by a heterologous nucleic acid sequence, thus inactivating the atonal-associated allele. The nucleic acids may be used in a method of treating an animal for hearing impairment, joint disease, balance disorders, abnormal cell proliferation, disease related to loss of a functional atonal-associated nucleic protein. They may particularly be used to treat an animal with deficiency in cerebellar granule neurons or their precursors, and also be used in promoting mechanoreceptive cell growth and generation of hair cells. The present sequence represents an atonal-associated acid sequence referred to in the invention. Note: The present sequence not shown in the specification, but was obtained from GenBank

SQ Sequence 354 AA;
 Query Match 100.0%; Score 110; DB 4; Length 354;
 Best Local Similarity 100.0%; Pred. No. 2e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AANARERRMHGHNHAFDQLR 21
 |||||

Db 163 AANARERRMHGHNHAFDQLR 183
 |||||

RESULT 12
 AAO30945
 ID AAO30945 standard; protein; 354 AA.
 AC AAO30945;
 XX
 DT 22-SEP-2003 (first entry)
 XX
 DE Human atonal homologue 1 (Hath1) protein #2.
 XX
 KW Atonal-associated protein; gastrointestinal; damaged intestinal tissue;
 cancer; inflammatory bowel disease; irritable bowel syndrome; infection;
 KW necrotising enterocolitis; hearing impairment; abnormal proliferation;
 joint disease; gene therapy; imbalance disorder; atonal homologue 1;
 KW Hath1; human.
 XX
 OS Homo sapiens.
 XX
 PN WO2003047532-A2.
 PD 12-JUN-2003.
 XX
 PF 03-DEC-2002; 2002WO-US041458.
 XX
 PR 05-DEC-2001; 2001US-00004717.
 XX
 PA (BAYU) BAYLOR COLLEGE MEDICINE.
 XX
 PI Zoghbi HY, Qi Y;
 XX
 DR WPI; 2003-505253/47.
 DR N-PSDB; AAL62230.
 XX
 PT New compositions comprising at least one stem cell which is up- or down-
 regulated for expression of an atonal-associated sequence, and at least
 PT one regulatory factor, useful for treating a gastrointestinal condition,
 PT e.g. cancer.
 XX
 PS Claim 33; Page 151-152; 157pp; English.
 XX
 CC The invention relates to pharmaceutical compositions and methods for the
 therapeutic use of an atonal-associated sequence. The composition
 CC comprises of a stem cell which is up- or down-regulated for expression of
 CC an atonal-associated sequence and a regulatory factor. Atonal-associated
 CC amino acid or nucleic acid sequence is used to treat a gastrointestinal
 CC condition including cancer, damaged intestinal tissue, inflammatory bowel
 CC disease, irritable bowel syndrome, infection or necrotising enterocolitis.
 CC They are also useful for promoting mechanoreceptive cell growth, for
 CC generating hair cells, for treating hearing impairment or imbalance
 CC disorder, joint disease, abnormal proliferation (e.g. cancer), or a
 CC disease that is a result or loss of functional atonal-associated
 CC sequences. Atonal-associated sequences are also used in gene therapy. The
 CC present sequence is human atonal homologue 1 (Hath1) protein. This
 CC sequence is used to illustrate the method of the invention
 XX
 SQ Sequence 354 AA;
 Query Match 100.0%; Score 110; DB 6; Length 354;
 Best Local Similarity 100.0%; Pred. No. 2e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AANARERRMHGHNHAFDQLR 21
 |||||

Db 163 AANARERRMHGHNHAFDQLR 183
 |||||

RESULT 14
 AAB60366
 ID AAB60366 standard; protein; 266 AA.

```

XX AC AAB60366;
XX DT 24-APR-2001 (first entry)
XX DE Zebrafish atonal homologue-1 (zath-1) protein, SEQ ID NO:40.
XX KW Atonal; homologue; orthologue; atonal-associated protein; deafness;
XX KW hearing impairment; vestibular effect; balance disorder; osteoarthritis;
XX KW cellular proliferation; cerebellar granule neuron; gene therapy;
XX KW mechanoreceptive cell growth; auditory; osteopathic; cytostatic;
XX KW transgenic animal.
XX OS Danio rerio.
XX PN WO200073764-A2.
XX PD 07-DEC-2000.
XX PF 01-JUN-2000; 2000WO-US015410.
XX PR 01-JUN-1999; 99US-0137060P.
XX PR 19-JAN-2000; 2000US-0176933P.
XX PA (BAYU ) BAYLOR COLLEGE MEDICINE.
XX PI Zoghbi HY, Bellen H, Birmingham N, Hassan B, Ben-Arie N;
XX PI WPI; 2001-0322190/04.
XX DR N-PSDB; AAF27274.
XX PT Therapeutic use of atonal-associated nucleic acids or amino acids, or any
XX PT of its homologs or orthologs, for the treatment of e.g. deafness,
XX PT osteoarthritis and abnormal cell proliferation.
XX PS Disclosure; Page; 142pp; English.
XX CC The invention relates to the use of atonal-associated nucleic acid or
XX CC amino acid sequence, or any of its homologues or orthologues as
XX CC therapeutic agents for the treatment of deafness, partial hearing loss,
XX CC vestibular effects due to damage or loss of inner hair cells,
XX CC osteoarthritis and abnormal cell proliferation. The invention also
XX CC encompasses methods of screening for compounds which affect the
XX CC expression of an atonal-associated nucleic acid sequence in an animal,
XX CC and a transgenic animal in which an allele of a native atonal-associated
XX CC gene is replaced by a heterologous nucleic acid sequence, thus
XX CC inactivating the atonal-associated allele. The nucleic acids or proteins
XX CC may be used in a method of treating an animal for hearing impairment,
XX CC joint disease, balance disorders, abnormal cell proliferation, or other
XX CC disease related to loss of a functional atonal-associated nucleic acid or
XX CC protein. They may particularly be used to treat an animal with a
XX CC deficiency in cerebellar granule neurons or their precursors, and may
XX CC also be used in promoting mechanoreceptive cell growth and generating
XX CC hair cells. The present sequence represents an atonal-associated amino
XX CC acid sequence referred to in the invention. Note: The present sequence is
XX CC not shown in the specification, but was obtained from GenBank
XX SQ Sequence 266 AA;
XX Query Match 97.3%; Score 107; DB 4; Length 266;
XX Best Local Similarity 95.2%; Pred. No. 4.9e-10;
XX Matches 20; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 AANARERRMHGHLNHFADQLR 21
DB 122 AANARERRMHGHLNHFADQLR 142
RESULT 15
ID AAO30936
XX AAO30936 standard; protein; 266 AA.
XX AAO30936;

```

```

XX DT 23-OCT-2003 (revised)
XX DT 22-SEP-2003 (first entry)
XX DE Zebra fish atonal homologue 1 protein #1.
XX KW Atonal-associated protein; gastrointestinal; damaged intestinal tissue;
XX KW cancer; inflammatory bowel disease; irritable bowel syndrome; infection;
XX KW necrotising enterocolitis; hearing impairment; abnormal proliferation;
XX KW joint disease; gene therapy; imbalance disorder; atonal homologue 1;
XX KW zebra fish.
XX OS Danio rerio.
XX PN WO2003047532-A2.
XX PD 12-JUN-2003.
XX PF 03-DEC-2002; 2002WO-US041458.
XX PR 05-DEC-2001; 2001US-00004717.
XX PA (BAYU ) BAYLOR COLLEGE MEDICINE.
XX PI Zoghbi HY, Qi Y;
XX PI WPI; 2003-505253/47.
XX DR N-PSDB; AAL62221.
XX PT New compositions comprising at least one stem cell which is up- or down-
XX PT regulated for expression of an atonal-associated sequence, and at least
XX PT one regulatory factor, useful for treating a gastrointestinal condition,
XX PT e.g. cancer.
XX PS Disclosure; Page 139-140; 157pp; English.
XX CC The invention relates to pharmaceutical compositions and methods for the
XX CC therapeutic use of an atonal-associated sequence. The composition
XX CC comprises of a stem cell which is up- or down-regulated for expression of
XX CC an atonal-associated sequence and a regulatory factor. Atonal-associated
XX CC amino acid or nucleic acid sequence is used to treat a gastrointestinal
XX CC condition including cancer, damaged intestinal tissue, inflammatory bowel
XX CC disease, irritable bowel syndrome, infection or necrotising enterocolitis.
XX CC They are also useful for promoting mechanoreceptive cell growth, for
XX CC generating hair cells, for treating hearing impairment or imbalance
XX CC disorder, joint disease, abnormal proliferation (e.g. cancer), or a
XX CC disease that is a result or loss of functional atonal-associated
XX CC sequences. Atonal-associated sequences are also used in gene therapy. The
XX CC present sequence is zebra fish atonal homologue 1 protein. This sequence
XX CC is used to illustrate the method of the invention. (Updated on 23-OCT-
XX CC 2003 to standardise OS field)
XX SQ Sequence 266 AA;
XX Query Match 97.3%; Score 107; DB 6; Length 266;
XX Best Local Similarity 95.2%; Pred. No. 4.9e-10;
XX Matches 20; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 AANARERRMHGHLNHFADQLR 21
DB 122 AANARERRMHGHLNHFADQLR 142
Search completed: September 21, 2004, 21:20:57
Job time : 10.912 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 21:18:48 ; Search time 2.576 Seconds
(without alignments)
420.864 Million cell updates/sec

Title: US-09-980-381A-70
Perfect score: 110
Sequence: 1 AANARERRMGLNHAFDQLR 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	57	4	US-08-722-570-8
2	110	100.0	57	4	US-08-932-411A-8
3	84	76.4	57	4	US-08-722-570-9
4	84	76.4	57	4	US-08-932-411A-9
5	81	73.6	57	4	US-08-722-570-6
6	81	73.6	57	4	US-08-932-411A-6
7	81	73.6	352	1	US-08-552-142A-4
8	81	73.6	352	1	US-08-910-973-4
9	81	73.6	352	4	US-09-499-227-4
10	81	73.6	352	5	PCT-US95-05741-4
11	80	72.7	57	4	US-08-722-570-5
12	80	72.7	57	4	US-08-932-411A-5
13	80	72.7	156	1	US-08-552-142A-9
14	80	72.7	156	1	US-08-910-973-9
15	80	72.7	156	4	US-09-499-227-9
16	80	72.7	156	5	PCT-US95-05741-9
17	80	72.7	356	1	US-08-552-142A-15
18	80	72.7	356	1	US-08-910-973-15
19	80	72.7	356	3	US-09-234-332-7
20	80	72.7	356	3	US-09-234-332-8
21	80	72.7	356	4	US-09-499-227-15
22	80	72.7	357	1	US-08-552-142A-2
23	80	72.7	357	1	US-08-910-973-2
24	80	72.7	357	4	US-09-499-227-2
25	80	72.7	357	5	PCT-US95-05741-2
26	73	66.4	57	4	US-08-722-570-7
27	73	66.4	57	4	US-08-932-411A-7

28 73 66.4 379 1 US-08-552-142A-11 Sequence 11, Appl
29 73 66.4 380 1 US-08-552-142A-17 Sequence 17, Appl
30 73 66.4 381 1 US-08-910-973-11 Sequence 11, Appl
31 73 66.4 381 4 US-09-499-227-11 Sequence 11, Appl
32 73 66.4 382 3 US-08-910-973-17 Sequence 17, Appl
33 73 66.4 382 3 US-09-234-332-9 Sequence 9, Appl
34 73 66.4 382 4 US-09-499-227-17 Sequence 17, Appl
35 69 62.7 57 4 US-08-722-570-4 Sequence 4, Appl
36 69 62.7 57 4 US-08-932-411A-4 Sequence 4, Appl
37 69 62.7 214 4 US-08-722-570-2 Sequence 2, Appl
38 69 62.7 214 4 US-08-932-411A-2 Sequence 2, Appl
39 68 61.8 214 4 US-08-932-411A-20 Sequence 20, Appl
40 67 60.9 57 4 US-08-722-570-3 Sequence 3, Appl
41 67 60.9 57 4 US-08-932-411A-3 Sequence 3, Appl
42 67 60.9 103 1 US-08-552-142A-13 Sequence 13, Appl
43 67 60.9 237 1 US-08-910-973-13 Sequence 13, Appl
44 67 60.9 237 4 US-09-499-227-13 Sequence 13, Appl
45 67 60.9 244 1 US-08-910-973-22 Sequence 22, Appl

ALIGNMENTS

RESULT 1
US-08-722-570-8
; Sequence 8, Application US/08722570
; Patent No. 6555337
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Ma, Glufu
; TITLE OF INVENTION: NEUROGENIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/722,570
; APPLICATION NUMBER: US/08/722,570
; FILING DATE: 27-SEP-1996
; CLASSIFICATION: 5365
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-63902/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-722-570-8

Query Match 100.0%; Score 110; DB 4; Length 57;
Best Local Similarity 100.0%; Pred. No. 1.8e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRMGLNHAFDQLR 21

Db 6 AANARERRMGLNHAFDQLR 26

```
RESULT 2
US-08-932-411A-8
; Sequence 8, Application US/08932411A
; Patent No. 6566496
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Ma, Qiufu
; TITLE OF INVENTION: NEUROGENIN
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr Hobbach Test Albritton & Herbert LLP
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/932,411A
; FILING DATE: 15-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/772,009
; FILING DATE: 19-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/722,570
; FILING DATE: 19-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-63902-3/RFT/RMS
; TELEPHONE: (415) 781-1989
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-932-411A-8
Query Match 100.0%; Score 110; DB 4; Length 57;
Best Local Similarity 100.0%; Pred. No. 1.8e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRMHGLNHAFDQLR 21
Db 6 AANARERRMHGLNHAFDQLR 26

RESULT 3
US-08-722-570-9
; Sequence 9, Application US/08722570
; Patent No. 6555337
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Ma, Qiufu
; TITLE OF INVENTION: NEUROGENIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
```

```
ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/722,570
; FILING DATE: 27-SEP-1996
; CLASSIFICATION: 536S
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-63902/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEX: (415) 398-3249
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-722-570-9
Query Match 76.4%; Score 84; DB 4; Length 57;
Best Local Similarity 81.0%; Pred. No. 3.8e-07;
Matches 17; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AANARERRMHGLNHAFDQLR 21
Db 6 AANARERRMHGLNHAFDQLR 26

RESULT 4
US-08-932-411A-9
; Sequence 9, Application US/08932411A
; Patent No. 6566496
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Ma, Qiufu
; TITLE OF INVENTION: NEUROGENIN
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr Hobbach Test Albritton & Herbert LLP
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/932,411A
; FILING DATE: 15-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/772,009
; FILING DATE: 19-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/722,570
; FILING DATE: 19-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-63902-3/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
```

TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 57 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-932-411A-9

Query Match 76.4%; Score 84; DB 4; Length 57;
Best Local Similarity 81.0%; Pred. No. 3.8e-07;
Matches 17; Conservative 1; Mismatches 3; Indels 0;

Qy 1 AANARERRMHGLNHAFDQLR 21
|||||
Db 6 AANARERRMHGLNHAFDQLR 26

RESULT 5
US-08-722-570-6

Sequence 6, Application US/08722570
Patent No. 655337
GENERAL INFORMATION:
APPLICANT: Anderson, David J.
APPLICANT: Ma, Qifu
TITLE OF INVENTION: NEUROGENIN
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr Hobbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/722,570
FILING DATE: 27-SEP-1996
CLASSIFICATION: 5365
ATTORNEY/AGENT INFORMATION:
NAME: Silva, Robin M.
REGISTRATION NUMBER: 38,304
REFERENCE/DOCKET NUMBER: A-63902/RFT/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 57 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-722-570-6

Query Match 73.6%; Score 81; DB 4; Length 57;
Best Local Similarity 80.0%; Pred. No. 1.2e-06;
Matches 16; Conservative 0; Mismatches 4; Indels 0;

Qy 2 AANARERRMHGLNHAFDQLR 21
|||||
Db 7 AANARERRMHGLNHAFDQLR 26

RESULT 6
US-08-932-411A-6

Sequence 6, Application US/08932411A
Patent No. 6566496
GENERAL INFORMATION:
APPLICANT: Anderson, David J.
APPLICANT: Ma, Qifu
TITLE OF INVENTION: NEUROGENIN
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr Hobbach Test Albritton & Herbert LLP
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/932,411A
FILING DATE: 15-SEP-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/772,009
FILING DATE: 19-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/722,570
FILING DATE: 19-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Silva, Robin M.
REGISTRATION NUMBER: 38,304
REFERENCE/DOCKET NUMBER: A-63902-3/RFT/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 57 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-932-411A-6

Query Match 73.6%; Score 81; DB 4; Length 57;
Best Local Similarity 80.0%; Pred. No. 1.2e-06;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 AANARERRMHGLNHAFDQLR 21
|||||
Db 7 AANARERRMHGLNHAFDQLR 26

RESULT 7

US-08-552-142A-4
Sequence 4, Application US/08552142A
Patent No. 5695995
GENERAL INFORMATION:
APPLICANT: Weintraub, Harold M.
APPLICANT: Lee, Jacqueline E.
APPLICANT: Tapscott, Stephen J.
APPLICANT: Hollenberg, Stanley M.
TITLE OF INVENTION: Neurogenic Differentiation (NeuroD) Genes
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Christensen O'Connor Johnson Kindness PLLC
STREET: 1420 Fifth Avenue, Suite 2800
CITY: Seattle
STATE: WA
COUNTRY: USA

ZIP: 98101-2347
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 02-NOV-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,238
FILING DATE: 06-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US95/05741
FILING DATE: 08-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Broderick, Thomas F.
REGISTRATION NUMBER: 31,332
REFERENCE/DOCKET NUMBER: FPCR-1-8933
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-682-8100
TELEFAX: 206-225-0709
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-552-142A-4

Query Match 73.6%; Score 81; DB 1; Length 352;
Best Local Similarity 80.0%; Pred. No. 9.6e-06;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ANARERRMGLNHAFQRLR 21
Db 107 ANARERNRMEGLNDALDSLRL 126

RESULT 8
US-08-910-973-4
Sequence 4, Application US/08910973
Patent No. 5795723
GENERAL INFORMATION:
APPLICANT: Tapscott, Stephen J.
APPLICANT: Olson, James M.
TITLE OF INVENTION: Expression of Neurogenic bHLH Genes in Primitive Neuroectoder
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Christensen O'Connor Johnson KindnessPLLC
STREET: 1420 Fifth Avenue, Suite 2800
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101-2347
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: US/08/910,973
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,238
FILING DATE: 06-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US95/05741
FILING DATE: 08-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/17532

FILING DATE: 30-October-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sheiness, Diana K.
REGISTRATION NUMBER: 35,356
REFERENCE/DOCKET NUMBER: FPCR-1-10958
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-682-8100; 206-224-0735 (direct)
TELEFAX: 206-225-0779
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-910-973-4
Query Match 73.6%; Score 81; DB 1; Length 352;
Best Local Similarity 80.0%; Pred. No. 9.6e-06;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 ANARERRMGLNHAFQRLR 21
Db 107 ANARERNRMEGLNDALDSLRL 126
RESULT 9
US-09-499-227-4
Sequence 4, Application US/09499227
Patent No. 644463
GENERAL INFORMATION:
APPLICANT: Tapscott, Stephen J.
APPLICANT: Olson, James M.
TITLE OF INVENTION: Expression of Neurogenic bHLH Genes in Primitive Neuroectoder
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Christensen O'Connor Johnson KindnessPLLC
STREET: 1420 Fifth Avenue, Suite 2800
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101-2347
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: US/09/499,227
APPLICATION NUMBER: US/09/499,227
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,238
FILING DATE: 06-May-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US95/05741
FILING DATE: 08-May-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/17532
FILING DATE: 30-October-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/910,973
FILING DATE: 07-August-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sheiness, Diana K.
REGISTRATION NUMBER: 35,356
REFERENCE/DOCKET NUMBER: FPCR-1-12742
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-682-8100; 206-224-0735 (direct)
TELEFAX: 206-225-0779
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: amino acid
TOPOLOGY: linear


```
; MOLECULE TYPE: protein
US-09-499-227-4
Query Match 73.6%; Score 81; DB 4; Length 352;
Best Local Similarity 80.0%; Pred. No. 9.6e-06;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ANARERRMHGLNHAFDQLR 21
||||| ||||| ||||| |||||
Db 107 ANARERNRMHGLNDALESLR 126

RESULT 10
PCT-US95-05741-4
; Sequence 4, Application PC/TUS9505741
; GENERAL INFORMATION:
; APPLICANT: Weintraub, Harold
; APPLICANT: Lee, Jacqueline E.
; APPLICANT: Tapscott, Stephen J.
; APPLICANT: Hollenberg, Stanley M.
; TITLE OF INVENTION: Neurogenic Differentiation (NeuroD) Gene
; TITLE OF INVENTION: and Protein
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen O'Connor Johnson Kindness
; STREET: 1420 Fifth Avenue, Suite 2800
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101-2347
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/05741
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Broderick, Thomas F.
; REGISTRATION NUMBER: 31,332
; REFERENCE/DOCKET NUMBER: FHCR-1-8504
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-682-8100
; TELEFAX: 206-225-0709
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-05741-4
Query Match 73.6%; Score 81; DB 5; Length 352;
Best Local Similarity 80.0%; Pred. No. 9.6e-06;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ANARERRMHGLNHAFDQLR 21
||||| ||||| ||||| |||||
Db 107 ANARERNRMHGLNDALESLR 126

RESULT 11
US-08-722-570-5
; Sequence 5, Application US/08722570
; Patent No. 6555337
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Ma, Qiufu
; TITLE OF INVENTION: NEUROGENIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr Hobbach Test Albritton & Herbert LLP
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/932,411A
; FILING DATE: 15-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/772,009
; FILING DATE: 19-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/722,570
; FILING DATE: 19-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; ADDRESS: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/722,570
; FILING DATE: 27-SEP-1996
; CLASSIFICATION: 5365
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-63902/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-722-570-5
Query Match 72.7%; Score 80; DB 4; Length 57;
Best Local Similarity 80.0%; Pred. No. 1.8e-06;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ANARERRMHGLNHAFDQLR 21
||||| ||||| ||||| |||||
Db 7 ANARERNRMHGLNDALESLR 26

RESULT 12
US-08-932-411A-5
; Sequence 5, Application US/08932411A
; Patent No. 6586496
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Ma, Qiufu
; TITLE OF INVENTION: NEUROGENIN
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr Hobbach Test Albritton & Herbert LLP
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/932,411A
; FILING DATE: 15-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/772,009
; FILING DATE: 19-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/722,570
; FILING DATE: 19-DEC-1996
; ATTORNEY/AGENT INFORMATION:
```

NAME: Silva, Robin M.
REGISTRATION NUMBER: 38,304
REFERENCE/DOCKET NUMBER: A-63902-3/RFT/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 57 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-932-411A-5

Query Match 72.7%; Score 80; DB 4; Length 57;
Best Local Similarity 80.0%; Pred. No. 1.8e-06;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ANARERRRHGHNHAFDQLR 21
||||| ||||| ||||| ||||| |||||
Db 7 ANARERNRHGHNALNLR 26

RESULT 13
US-08-552-142A-9
Sequence 9, Application US/08552142A
Patent No. 5695995
GENERAL INFORMATION:
APPLICANT: Weintraub, Harold M.
APPLICANT: Lee, Jacqueline E.
APPLICANT: Tapscott, Stephen J.
APPLICANT: Hollenberg, Stanley M.
TITLE OF INVENTION: Neurogenic Differentiation (Neurod) Genes
TITLE OF INVENTION: And Proteins
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Christensen O'Connor Johnson KindnessPLLC
STREET: 1420 Fifth Avenue, Suite 2800
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101-2347
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 02-NOV-1995
APPLICATION NUMBER: US/08/552,142A
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,238
FILING DATE: 06-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US95/05741
FILING DATE: 08-MAY-1995
NAME: Broderick, Thomas F.
REGISTRATION NUMBER: 31,332
REFERENCE/DOCKET NUMBER: FHCR-1-8933
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-682-8100
TELEFAX: 206-225-0709
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 156 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-552-142A-9

Query Match 72.7%; Score 80; DB 1; Length 156;
Best Local Similarity 80.0%; Pred. No. 5.6e-06;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ANARERRRHGHNHAFDQLR 21
||||| ||||| ||||| ||||| |||||
Db 106 ANARERNRHGHNALNLR 125

RESULT 14
US-08-910-973-9
Sequence 9, Application US/08910973
Patent No. 5795723
GENERAL INFORMATION:
APPLICANT: Tapscott, Stephen J.
APPLICANT: Olson, James M.
TITLE OF INVENTION: Expression of Neurogenic bHLH Genes in Primitive Neuroectoder
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Christensen O'Connor Johnson KindnessPLLC
STREET: 1420 Fifth Avenue, Suite 2800
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101-2347
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: US/08/910,973
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,238
FILING DATE: 06-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US95/05741
FILING DATE: 08-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/17532
FILING DATE: 30-October-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sheiness, Diana K.
REGISTRATION NUMBER: 35,356
REFERENCE/DOCKET NUMBER: FHCR-1-10958
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-682-8100; 206-224-0735 (direct)
TELEFAX: 206-225-0779
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 156 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-910-973-9

Query Match 72.7%; Score 80; DB 1; Length 156;
Best Local Similarity 80.0%; Pred. No. 5.6e-06;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ANARERRRHGHNHAFDQLR 21
||||| ||||| ||||| ||||| |||||
Db 106 ANARERNRHGHNALNLR 125

RESULT 15
US-09-499-227-9
Sequence 9, Application US/09499227
Patent No. 644463
GENERAL INFORMATION:

APPLICANT: Tapscott, Stephen J.
APPLICANT: Olson, James M.
TITLE OF INVENTION: Expression of Neurogenic bHLH Genes in Primitive Neuroectoder
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Christensen O'Connor Johnson Kindness PLLC
STREET: 1420 Fifth Avenue, Suite 2800
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101-2347
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/499,227
FILING DATE: 05-August-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,238
FILING DATE: 06-May-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US95/05741
FILING DATE: 08-May-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/17532
FILING DATE: 30-October-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/910,973
FILING DATE: 07-August-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sheiness, Diana K.
REGISTRATION NUMBER: 35,356
REFERENCE/DOCKET NUMBER: FHC9-1-12742
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-682-8100; 206-224-0735 (direct)
TELEFAX: 206-225-0779
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 156 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-499-227-9

Query Match 72.7%; Score 80; DB 4; Length 156;
Best Local Similarity 80.0%; Pred. No. 5.6e-06;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ANARERRMHGLNHAFDQLR 21
DB 106 ANARERRMHGLNHALDNLK 125

Search completed: September 21, 2004, 21:25:51
Job time : 2.576 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 21:24:21 ; Search time 9.576 Seconds
(without alignments)
704.240 Million cell updates/sec

Title: US-09-980-381A-70

Perfect score: 110
Sequence: 1 AANARERRMHGLNHAFDQLR 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1342398 seqs, 321133274 residues

Total number of hits satisfying chosen parameters: 1342398

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	110	100.0	21	16	US-10-373-249-5
2	110	100.0	57	8	US-08-722-570-8
3	110	100.0	57	12	US-10-425-259-8
4	110	100.0	161	13	US-10-004-717-60
5	110	100.0	259	13	US-10-004-717-66
6	110	100.0	351	13	US-10-004-717-11
7	110	100.0	351	13	US-10-004-717-46
8	110	100.0	351	16	US-10-373-249-1
9	110	100.0	351	16	US-10-373-249-2
10	110	100.0	354	13	US-10-004-717-2
11	110	100.0	354	13	US-10-004-717-58
12	110	100.0	354	16	US-10-373-249-3
13	110	100.0	354	16	US-10-373-249-4
14	107	97.3	266	13	US-10-004-717-40
15	98	89.1	46	13	US-10-004-717-62

16	91	82.7	45	13	US-10-004-717-54	Sequence 54, Appl
17	88	80.0	149	13	US-10-004-717-23	Sequence 23, Appl
18	88	80.0	151	13	US-10-004-717-36	Sequence 36, Appl
19	87	79.1	138	13	US-10-004-717-50	Sequence 50, Appl
20	87	79.1	138	13	US-10-004-717-52	Sequence 52, Appl
21	84	76.4	45	13	US-10-004-717-56	Sequence 56, Appl
22	84	76.4	57	8	US-08-722-570-9	Sequence 9, Appl
23	84	76.4	57	12	US-10-425-259-9	Sequence 9, Appl
24	84	76.4	312	13	US-10-004-717-64	Sequence 64, Appl
25	83	75.5	325	15	US-10-120-801-96	Sequence 96, Appl
26	82	74.5	189	13	US-10-004-717-17	Sequence 17, Appl
27	81	73.6	57	8	US-08-722-570-6	Sequence 6, Appl
28	81	73.6	57	12	US-10-425-259-6	Sequence 6, Appl
29	81	73.6	337	12	US-10-221-625-20	Sequence 20, Appl
30	81	73.6	337	13	US-10-004-717-9	Sequence 9, Appl
31	81	73.6	337	13	US-10-004-717-44	Sequence 44, Appl
32	81	73.6	352	16	US-10-654-102-6	Sequence 6, Appl
33	81	73.6	352	16	US-10-654-102-16	Sequence 16, Appl
34	81	73.6	352	16	US-10-654-102-27	Sequence 27, Appl
35	81	73.6	352	16	US-10-654-102-38	Sequence 38, Appl
36	81	73.6	352	16	US-10-654-102-43	Sequence 43, Appl
37	80	72.7	57	8	US-08-722-570-5	Sequence 5, Appl
38	80	72.7	57	12	US-10-425-259-5	Sequence 5, Appl
39	80	72.7	113	16	US-10-654-102-30	Sequence 30, Appl
40	80	72.7	113	16	US-10-654-102-41	Sequence 41, Appl
41	80	72.7	156	16	US-10-654-102-37	Sequence 37, Appl
42	80	72.7	216	16	US-10-654-102-23	Sequence 23, Appl
43	80	72.7	285	16	US-10-654-102-29	Sequence 29, Appl
44	80	72.7	330	13	US-10-004-717-33	Sequence 33, Appl
45	80	72.7	330	13	US-10-004-717-48	Sequence 48, Appl

ALIGNMENTS

RESULT 1

US-10-373-249-5
; Sequence 5, Application US/10373249
; Publication No. US20040166091A1
; GENERAL INFORMATION:
; APPLICANT: BROUGH, DOUGLAS E
; TITLE OF INVENTION: MATERIALS AND METHODS FOR TREATING DISORDERS OF THE EAR
; FILE REFERENCE: 214680
; CURRENT APPLICATION NUMBER: US/10/373,249
; CURRENT FILING DATE: 2003-02-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent version 3.1
; SEQ ID NO 5
; LENGTH: 21
; TYPE: PRT
; ORGANISM: drosophila
US-10-373-249-5

Query Match 100.0%; Score 110; DB 16; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRMHGLNHAFDQLR 21
DB 1 AANARERRMHGLNHAFDQLR 21

RESULT 2

US-08-722-570-8
; Sequence 8, Application US/08722570
; Publication No. US20030044887A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Ma, Oufu
; TITLE OF INVENTION: NEUROGENIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohnbach, Test, Albritton & Herbert

Wed Sep 22 12:21:31 2004

```
; CURRENT APPLICATION NUMBER: US/10/004,717
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 09/585,645
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: 60/176,993
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: 60/137,060
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-004-717-46

Query Match      100.0%; Score 110; DB 13; Length 351;
Best Local Similarity 100.0%; Pred. No. 8e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRRMHGLNHAFDQLR 21
   ||||||||||||||||||||
Db 160 AANARERRRMHGLNHAFDQLR 180

RESULT 6
US-10-373-249-1
; Sequence 1, Application US/10373249
; Publication No. US20040166091A1
; GENERAL INFORMATION:
; APPLICANT: BROUGH, DOUGLAS E
; TITLE OF INVENTION: MATERIALS AND METHODS FOR TREATING DISORDERS OF THE EAR
; FILE REFERENCE: 214680
; CURRENT APPLICATION NUMBER: US/10/373,249
; CURRENT FILING DATE: 2003-02-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-373-249-1

Query Match      100.0%; Score 110; DB 16; Length 351;
Best Local Similarity 100.0%; Pred. No. 8e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRRMHGLNHAFDQLR 21
   ||||||||||||||||||||
Db 160 AANARERRRMHGLNHAFDQLR 180

RESULT 9
US-10-373-249-2
; Sequence 2, Application US/10373249
; Publication No. US20040166091A1
; GENERAL INFORMATION:
; APPLICANT: BROUGH, DOUGLAS E
; TITLE OF INVENTION: MATERIALS AND METHODS FOR TREATING DISORDERS OF THE EAR
; FILE REFERENCE: 214680
; CURRENT APPLICATION NUMBER: US/10/373,249
; CURRENT FILING DATE: 2003-02-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-373-249-2

Query Match      100.0%; Score 110; DB 16; Length 351;
Best Local Similarity 100.0%; Pred. No. 8e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRRMHGLNHAFDQLR 21
   ||||||||||||||||||||
Db 160 AANARERRRMHGLNHAFDQLR 180

; CURRENT APPLICATION NUMBER: US/10/004,717
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 09/585,645
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: 60/176,993
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: 60/137,060
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 66
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Frog
US-10-004-717-66

Query Match      100.0%; Score 110; DB 13; Length 259;
Best Local Similarity 100.0%; Pred. No. 5.8e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRRMHGLNHAFDQLR 21
   ||||||||||||||||||||
Db 115 AANARERRRMHGLNHAFDQLR 135

RESULT 6
US-10-004-717-11
; Sequence 11, Application US/10004717
; Publication No. US20020192665A1
; GENERAL INFORMATION:
; APPLICANT: ZOGHBI, HUDA Y.
; APPLICANT: YANG, QI
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC USE OF AN
; TITLE OF INVENTION: ATONAL ASSOCIATED SEQUENCE FOR DEAFNESS,
; TITLE OF INVENTION: OSTEOARTHRITIS, AND ABNORMAL CELL PROLIFERATION
; FILE REFERENCE: P01899US4
; CURRENT APPLICATION NUMBER: US/10/004,717
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 09/585,645
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: 60/176,993
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: 60/137,060
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-004-717-11

Query Match      100.0%; Score 110; DB 13; Length 351;
Best Local Similarity 100.0%; Pred. No. 8e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRRMHGLNHAFDQLR 21
   ||||||||||||||||||||
Db 160 AANARERRRMHGLNHAFDQLR 180

RESULT 7
US-10-004-717-46
; Sequence 46, Application US/10004717
; Publication No. US20020192665A1
; GENERAL INFORMATION:
; APPLICANT: ZOGHBI, HUDA Y.
; APPLICANT: YANG, QI
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC USE OF AN
; TITLE OF INVENTION: ATONAL ASSOCIATED SEQUENCE FOR DEAFNESS,
; TITLE OF INVENTION: OSTEOARTHRITIS, AND ABNORMAL CELL PROLIFERATION
; FILE REFERENCE: P01899US4
; CURRENT APPLICATION NUMBER: US/10/004,717
```



```
; PRIOR APPLICATION NUMBER: 60/176,993
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: 60/137,060
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Zebra Fish
US-10-004-717-40

Query Match          97.3%; Score 107; DB 13; Length 266;
Best Local Similarity 95.2%; Pred.No. 1.7e-08;
Matches 20; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 AANARERRRMHGLNHAFDQLR 21
      |||||
Db      122 AANARERRRMHGLNHAFDELR 142

RESULT 15
US-10-004-717-62
; Sequence 62, Application US/10004717
; Publication No. US20020192665A1
; GENERAL INFORMATION:
; APPLICANT: ZOGHBI, HUDA Y.
; APPLICANT: YANG, QI
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC USE OF AN
; TITLE OF INVENTION: ATONAL ASSOCIATED SEQUENCE FOR DEAFNESS,
; TITLE OF INVENTION: OSTEOARTHRITIS, AND ABNORMAL CELL PROLIFERATION
; FILE REFERENCE: P01899US4
; CURRENT APPLICATION NUMBER: US/10/004,717
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 09/585,645
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: 60/176,993
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: 60/137,060
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Puffer Fish
US-10-004-717-62

Query Match          89.1%; Score 98; DB 13; Length 46;
Best Local Similarity 90.5%; Pred.No. 7e-08;
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 AANARERRRMHGLNHAFDQLR 21
      |||||
Db      1 AANARERRRMHGLNKAFDELR 21

Search completed: September 21, 2004, 21:40:26
Job time : 9.576 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 21:08:17 ; Search time 47.2 Seconds
(without alignments)
721.436 Million cell updates/sec

Title: US-09-980-381A-58
Perfect score: 1862
Sequence: 1 MSRLHAEEWAELVKGDLHH.....HRSDGFSPHSHYSDSEAS 354

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1627.5	87.4	351	2 A56387	helix-loop-helix t
2	219.5	11.8	347	2 JC7999	basic helix-loop-h
3	214	11.5	352	2 I51687	neurogenic differe
4	213	11.4	312	2 A40708	basic-helix-loop-h
5	210.5	11.3	381	2 G02668	neurogenic basic-h
6	207.5	11.1	381	2 A57059	beta-cell E-box tr
7	205.5	11.0	337	2 I48682	NEX-1 - mouse
8	205.5	11.0	337	2 I57038	basic helix-loop-h
9	202	10.8	357	2 JC4703	neurogenic differe
10	202	10.8	357	2 I49338	neuro D-related fa
11	202	10.8	383	2 JC4688	hypothetical prote
12	200	10.7	192	2 T15764	KW8 protein - rat
13	199	10.7	381	2 JC4647	transcription fact
14	195	10.5	311	2 S20085	helix-loop-helix t
15	195	10.5	367	2 JC6087	LIN-32 protein - C
16	184	9.9	71	2 T29378	T-cell acute lymph
17	179	9.6	331	2 A36358	SC1 protein homolo
18	178	9.6	329	2 A37864	lyl-1 protein - hu
19	175.5	9.4	267	2 A30988	transcription fact
20	174.5	9.4	326	2 S71755	helix-loop-helix p
21	174	9.3	278	2 A43814	lyl-1 protein - mo
22	174	9.3	278	2 S16678	protein kinase C-b
23	171	9.2	334	2 G02409	scleraxis - mouse
24	167	9.0	207	2 I53154	helix-loop-helix p
25	162.5	8.7	245	2 B43814	transcription fact
26	157.5	8.5	198	2 A57717	hypothetical prote
27	153.5	8.2	237	2 T15548	hypothetical prote
28	152	8.2	147	2 T29241	hypothetical prote
29	148.5	8.0	170	2 T29971	hypothetical prote

30	148	7.9	260	2 S11562	probable MASH-1 pr
31	147.5	7.9	360	2 A56066	basic helix-loop-h
32	146.5	7.9	206	2 I53066	gene M-twist prote
33	146	7.8	318	2 JC1171	muscle regulatory
34	145	7.8	318	2 A29636	myoblast determina
35	144.5	7.8	166	2 A33637	Xtwi protein - Afr
36	143	7.7	201	2 G01204	twist protein homo
37	140	7.5	319	2 S26827	myogenic factor 3
38	140	7.5	490	2 S00995	gene twist protein
39	136	7.3	319	2 S20086	MyoD1 protein - sh
40	136	7.3	382	2 B88561	protein F58A4.7b (
41	135.5	7.3	111	2 B41629	TAL2 protein - mou
42	135	7.3	345	2 B43731	achaeete-scute comp
43	134.5	7.2	292	2 S40979	hypothetical prote
44	134.5	7.2	1180	2 S69205	stripe a/b protein
45	134	7.2	231	2 S28186	achaeete-scute locu

ALIGNMENTS

RESULT 1

A56387
helix-loop-helix transcription factor MATH-1 - mouse
N:Alternate names: atonal homolog MATH-1
C:Species: Mus musculus (house mouse)
C>Date: 21-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 05-Nov-1999
C:Accession: A56387
R:Akazawa, C.; Ishibashi, M.; Shimizu, C.; Nakanishi, S.; Kageyama, R.
J. Biol. Chem. 270, 8730-8738, 1995
A:Title: A mammalian helix-loop-helix factor structurally related to the product of Dros
A:Reference number: A56387; MUID:95238366; PMID:7721778
A:Accession: A56387
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-351 <AKA>
A:Cross-references: GB:D43694; NID:g994770; PIDN:BAA07791.1; PID:di008377; PID:g994771
C:Genetics:
A:Introns: #status absent
C:Keywords: transcription factor

Query Match	87.4%;	Score	1627.5;	DB 2;	Length	351;			
Best Local Similarity	88.4%;	Pred. No.	5.7e-104;						
Matches	313;	Conservative	11;	Mismatches	27;	Indels	3;	Gaps	2;
QY	1	MSRLHAEEWAELVKGDLHHRQPHHLQPPPPPPPPATLQAREHPVYPPELSILLSTD	60						
Db	1	MSRLHAEEWAELVKGDLHHRHQP HHV--PELTPQPPATLQARDLFPVPAELSLSTD	58						
QY	61	PRAWLPTLOGICTARAAQYLLHSPELGASEAAAPRDEVDGRGLVRRSGGASSKSPG	120						
Db	59	PRAWLPTLOGICTARAAQYLLHSPELGASEAAAPRDEADDSQGLVRR--SGCGLSKSPG	117						
QY	121	PVKVRELQCLKGGVVVDDELGCSRQAPSSKQVNGVQKQRLAANARERRRMHGLNAFD	180						
Db	118	PVKVRELQCLKGGVVVDDELGCSRQAPSSKQVNGVQKQRLAANARERRRMHGLNAFD	177						
QY	181	QLRNVPISFNNDKKLSKYETLQMAQIYINALSELLQTPSGGEQPPPPASCKSDHHHLRT	240						
Db	178	QLRNVPISFNNDKKLSKYETLQMAQIYINALSELLQTPNVGEQPPPTASCKNDHHHLRT	237						
QY	241	AASVEGAGNATAAGAAQQAQSGGSGRTPPGSCRTFSAPASAGYSVQLDALHFSFTEDS	300						
Db	238	ASSYEGAGASAVAGA PPGGGPRTPPGCTRFSGPSSGGYSVQLDALHFPAFEDR	297						
QY	301	ALTAMQAQKLSPLPGSILQPVQEENSKTSPRSHRSDGFEFSPHSHYSDSDEAS	354						
Db	298	ALTAMQAQKLSPLPGSILQPVQEDNSKTSPRSHRSDGFEFSPHSHYSDSDEAS	351						

RESULT 2

IC7999

RESULT 2
JC7999

basic helix-loop-helix transcription factor, Zath3 - zebra fish
 C:Species: Brachydanio rerio (zebra fish)
 C:Date: 10-Nov-2003 #sequence_revision 10-Nov-2003 #text_change 24-Nov-2003
 C:Accession: J07999
 R:Pack, S.H.; Yeo, S.Y.; Yoo, K.W.; Hong, S.K.; Lee, S.; Rhee, M.; Chitnis, A.B.; Kim, C.
 Biochem. Biophys. Res. Commun. 308, 184-190, 2003
 A:Title: Zath3, a neural basic helix-loop-helix gene, regulates early neurogenesis in the zebra fish.
 A:Reference number: J07999; PMID:12890499
 A:Accession: J07999
 A:Molecule type: mRNA
 A:Residues: 1-347 <PAR>
 A:Cross-references: GB:AF204240
 A:Experimental source: (Embryo)
 C:Comment: This protein, which belongs to the family of basic helix-loop-helix transcriptional ganglia, and the roles in early neurogenesis in zebra fish.
 C:Genetics:
 A:Gene: zath3
 C:Keywords: basic helix-loop-helix protein; neurogenesis

Query Match 11.8%; Score 219.5; DB 2; Length 347;
 Best Local Similarity 24.6%; Pred. No. 7.8e-08;
 Matches 81; Conservative 41; Mismatches 112; Indels 95; Gaps 12;

QY 64 WLAPTL---QGICTARAAQYLLHSPGLGASAAAAPRDEVDGRLVRRSSGGASSKSPG 120
 Db 20 WMDELLSSQDGRTPRIGHYSLHRSNRGPLEIGS--EDMDEEEEEEDEEMGLDGERAP- 76
 QY 121 PVKVRQLCKLKGWVVDLGCQRAPSSKQVNGVOKO---RRLAANARERERMHGLN 176
 Db 77 -----KRGPKKKMTKAKQERPRARIKANARERSMHGLN 113
 QY 177 HAPQLNRVPSFNNDKKLYETLQMAQIYNALSELLQTPSGGEQF----- 224
 Db 114 DALDNLRRVMPVCSYKTKLKIETLRLARNYIWLSEVLES---GQSPESHGFVEMLCCKG 170
 QY 225 -PPPPASCKSDHHHLRTAAAYE-----GGAGNATAAGAAQASGGSQRTPPGSCRTFSA 278
 Db 171 LPQTSNLVAGCQLQGLPTTMLKDEKGVGAGVPGQGHPISYSPGLPSP----- 221
 QY 279 PASAGGYSVQLDALHFTFSDSALTAMMAQK-----NLSPS-----LPGSLQPVQE 325
 Db 222 -----YCTMAASHLLHLKYGKPPYENSSPNECSSGTPPYDGLPTPLSI 266
 QY 326 ENS---KTSRPSHRSDGEFSPH-SHYSDS 350
 Db 267 SGNFALKQEPSPREAERNFTPHPTIYSS 295

RESULT 3
 151687
 neurogenic differentiation factor - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
 C:Accession: I51687
 R:Lee, J.E.; Hollenberg, S.M.; Snider, L.; Turner, D.L.; Lipnick, N.; Weintraub, H.
 Science 268, 836-844, 1995
 A:Title: Conversion of Xenopus ectoderm into neurons by NeuroD, a basic helix-loop-helix protein.
 A:Reference number: A56481; MUID:95273957; PMID:7754368
 A:Accession: I51687
 A>Status: preliminary;
 A:Molecule type: mRNA
 A:Residues: 1-352 <LEE>
 A:Cross-references: EMBL:U28067; NID:g854740; PIDN:AAC59675.1; PID:g854741
 C:Genetics:
 A:Gene: neuroD

Query Match 11.5%; Score 214; DB 2; Length 352;
 Best Local Similarity 29.5%; Pred. No. 1.9e-07;
 Matches 70; Conservative 31; Mismatches 86; Indels 50; Gaps 9;

QY 144 RQAPSSKQVNGVQ-----KQRRLAANARERERMHGLNHFADQLNRNIPSFNNDKKLSKYE 199
 Db 83 KRRGPKKKMTKARVERFVRMKANARENRMHGLNDALDSLRKVVPCYSKTQKLSKIE 142

QY 200 TLQMAQIYNALSELL-----QTPSGGEQPPPP-----PASCK 232
 Db 143 TRLAKNYIWLSELLRSGKSPDLVSFVQTLCKGLSQDTNLVAGCLOLNPRTELPEQSQ 202
 QY 233 SDHHLRTAAAYEGAGNATAAGAAQASGGSQRTPPP-GSCRTFRSAPASAGGYSVQLDA 291
 Db 203 DIQSHMQTASS-----SFPLQGYQSPG--LSPPPYGTWDSHVFHVKPHSVGAALPE 254
 QY 292 LHFSTFEDSALTAMMAQKNLSPSLPGSIQPVQENSKTSPRSHRS-DGEFSPHSHY 347
 Db 255 F----FDSSTVT-----ECTSPSPDGLSPPLSVNGNFTFKHEHSEYDKNTFTTMY 302

RESULT 4

A40708
 basic-helix-loop-helix protein ato - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
 C:Accession: A40708
 R:Jarman, A.P.; Grau, Y.; Jan, L.Y.; Jan, Y.N.
 Cell 73, 1307-1321, 1993
 A:Title: atonal is a proneural gene that directs chordotonal organ formation in the Drosophila embryo.
 A:Reference number: A40708; MUID:93313961; PMID:8324823
 A:Accession: A40708
 A>Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-312 <JAR>
 A:Cross-references: GB:L36646; NID:g551565; PIDN:AAA21879.1; PID:g551566
 A:Note: sequence extracted from NCBI backbone (NCBIN:135094, NCBI:P:135095)
 C:Genetics:
 A:Gene: FlyBase:ato
 A:Cross-references: FlyBase:FBgn0010433

Query Match 11.4%; Score 213; DB 2; Length 312;
 Best Local Similarity 34.0%; Pred. No. 1.9e-07;
 Matches 68; Conservative 21; Mismatches 85; Indels 26; Gaps 6;

QY 33 PPQPRA-----TLQAREHPYVPPLSLDSDTP--RAWLAPTLOGICTARA 77
 Db 123 FVAPPPAVEVWGSNVGTCTIPASAAP--KPKRSYTKKNQPTTATSTPTAAAESASV 180
 QY 78 AQYLLHSPGLGASAAAAPRDEVDGRLVRRSSGGASSKSPGVKVRQLCKLKGWVV 137
 Db 181 NLYTEFFQNFDFDNSALFDDSDVEDDEDLN-LFSGGDFDQNGDSF-----DLADGENQ 232
 QY 138 DELGCSRQAPSSKQVNG-VQKORRLAANARERERMHGLNHFADQLNRNIPSFNNDKKLS 196
 Db 233 DAAAGSGKRGKQITPVVKKRRLAANARERERMQNLNQAFLRLRQYLPCLGNDRQLS 292
 QY 197 KYETLQMAQIYNALSELLQ 216
 Db 293 KHETLQMAQIYNALGDLRL 312

RESULT 5

G02668
 neurogenic basic-helix-loop-helix (bHLH) protein - human
 C:Species: Homo sapiens (man)
 C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 17-Jul-1998
 C:Accession: G02668
 R:Tapscott, S.J.; Tamimi, R.T.; McCormick, B.M.
 submitted to the EMBL Data Library, May 1996
 A:Reference number: H01573
 A:Accession: G02668
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-381 <TAP>
 A:Cross-references: EMBL:U58681; NID:g1477748; PID:g1477749
 C:Genetics:
 A:Gene: NeuroD2

Query Match 11.3%; Score 210.5; DB 2; Length 381;

[illegible]

RESULT 6
A57059
beta-cell E-box transcription activator 2 - hamster (fragment)
C:Species: Cricetinae gen. sp. (hamster)
C:Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 08-Oct-1999
C:Accession: A57059
R:Naya, F.J.; Stellrecht, C.M.M.; Tsai, M.J.
Genes Dev. 9, 1009-1019, 1995
A:Title: Tissue-specific regulation of the insulin gene by a novel basic helix-loop-helix
A:Reference number: A57059; MUID:95293222; PMID:7774807
A:Accession: A57059
A:Status: preliminary
A:Molecule_type: mRNA
A:Residues: 1-381 <NAY>

Query Match	11.1%;	Score	207.5;	DB	2;	Length	381;
Best Local Similarity	27.1%;	Pred. No.	5.7e-07;				
Matches	79;	Conservative	28;	Mismatches	66;	Indels	119;
						Gaps	12;
QY	144	QORAPSSKQVNGVQ----	KORLAA	ANARERRRMHGLN	HAFDQLRN	VPISFN	DNDKLSKYE
		::: :::	::: :::	::: :::	::: :::	::: :::	::: :::
Db	107	KRGGPKKKMTKARLER	PKLR	MRKANARERRMHGLN	AAALDNL	RKWP	CYSKTOKLSKIE
		::: :::	::: :::	::: :::	::: :::	::: :::	::: :::
QY	200	TLQMAQIYNALSELL----	-----	-----	-----	-----	-----
		::: :::	::: :::	::: :::	::: :::	::: :::	::: :::
Db	167	TLRLAKNYIWALESILL	RSGK	PDVSVFQTLCKG	LSQPTNL	IVAGCL	OLNRPFLPEQNP
		::: :::	::: :::	::: :::	::: :::	::: :::	::: :::
QY	225	PPPPASCKSDHHHLRTAA-----	-----	-----	-----	-----	-----
		::: :::	::: :::	::: :::	::: :::	::: :::	::: :::
Db	227	DMPP-----	HLTAS	AFVHPVYSQSG	LPSPPYGT	WDSHVQ	KVPPPHA-----
		::: :::	::: :::	::: :::	::: :::	::: :::	::: :::
QY	277	SAPASAGGYVOLDALH	FTFSD	SALTAMMAQK	NLSP	LPGLSQ	ILQVQ-----
		::: :::	::: :::	::: :::	::: :::	::: :::	::: :::
Db	274	-----	YSATLE-----	PPFES	PLTDC-----	TSP	SFDFGLSP
		::: :::	::: :::	::: :::	::: :::	::: :::	::: :::
QY	325	-----	EEN-----	SKTS	PRSHRS-----	DG-----	EFSPHSV
		::: :::	::: :::	::: :::	::: :::	::: :::	::: :::
Db	316	ARFENKVAFTMHP	PAATLAG	POSHGSI	FTSGAT	PRCEI	PDIMNMF
		::: :::	::: :::	::: :::	::: :::	::: :::	::: :::

RESULT 7
148682
NEX-1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 20-Jun-2000
C:Accession: I48682; S69346
R:Bartholoma, A.; Nave, K.A.
Mech.Dev. 48, 217-228, 1994
A:Title: NEX-1: a novel, brain-specific helix-loop-helix protein with autoregulation and
A:Reference number: I48682; MUID:95200803; PMID:7545978
A:Accession: I48682
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-337 <RES>
A:Cross-references: EMBL:U29086; NID:g881961; PIDN:AA014576.1; PID:g881962
R:Shimizu, C.; Akazawa, C.; Nakanishi, S.; Kageyama, R.
Eur. J. Biochem. 229, 239-248, 1995
A:Title: MATH-2, a mammalian helix-loop-helix factor structurally related to the product
A:Reference number: S69346; MUID:95262673; PMID:7744035
A:Accession: S69346
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-337 <SHI>
A:Cross-references: EMBL:D44480; NID:g994772; PIDN:BA007923.1; PID:g994773
C:Genetics:
A:Gene: nex-1

Query Match	11.0%	Score	205.5;	DB	2;	Length	337;
Best Local Similarity	29.0%;	Pred.	No. 6.8e-07;				
Matches	71;	Conservative	31;	Mismatches	94;	Indels	49; Gaps
							7;
QY	102	RGEIVRRSSGCASSSKPGPPVKVREQLCKLKGVVVDDELGCSRORAPSSKQWNGVQ-----	157				
		::: :	::: :				
DB	44	RGKIKRAGPGETEKEEEEEDREE-----DENGLSRRGLRKKTTLRLERV	92				
		:	:				
QY	158	KORLIANAREBRMRHGLNHAFDQLRNVIPSFNNDKKLSKVETLQMAQIYINAISELLOT	217				
		:	:				
DB	93	KFRSQEANARENRMHGLNLDNLNRKVPVCYSTQTKLSTKETLRKAKNYIWAUSEILRI	152				
		:	:				
QY	218	PSGREQP-----PPPPASCSKDHHHLRTAASVEGGGAGNATAAGAQAQGSGSQ	264				
		:	:				
DB	153	---GKRPLLTFVNCLKGLSQPTTNLVAGCLQL-NARSLFMQOGGEAAHHTRSYPSTFY	208				
		:	:				
QY	265	RP-----TPPGSCRFPSAPASGGYSVQLDALHPFDSDSALTAMWAQNKLSPSLPG	317				
		:	:				
DB	209	PPTHSEPLATPPGHGTLDNSKMFPNYICSAVESFYESTSPEC-----SPQEG	258				
		:	:				
QY	318	SILQP	322				
	:						
DB	259	PLSP	263				

RESULT 8
157038
gene Dlx-3 protein - mouse
C.Species: Mus sp. (mouse)
C.Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 28-Feb-1997
C.Accession: I57038
R.Robinson, G.W.; Mahon, K.A.
Mech. Dev. 48, 199-215, 1994
A.Title: Differential and overlapping expression domains of Dlx-2 and Dlx-3 suggest distinct
A.Reference number: I57038; MUID:95200802; PMID:7893603
A.Accession: I57038
A.Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: mRNA
A.Residues: 1-337 <RES>
A.Cross-references: GB:S75299; NID:gb986199; PID:gb986200
C.Genetics:
A.Gene: Dlx-3

Query Match	11.0%;	Score 205.5;	DB 2;	Length 337;
Best Local Similarity	29.0%;	Pred. No. 6.8e-07;		
Matches 71;	Conservative 31;	Mismatches 94;	Indels 49;	Gaps 7

Db 70 TLAEV-----KEGELGEEEEEE-EGUDE-----AEGRPKKRGPKKRWMT 113
 QY 128 LCKLKGVVVDDELGCSQRAPSSQVNGVQKORLIANARERRMHGHNHAFDQLRNVIP 187
 Db 114 KARLER-----SKLRQKANARERNRMDLNAALDNLKRVVP 150
 QY 188 SFNNDKLSKYETLQMAQIYNALSELL-----OTPSGGEOPPPP--PASC-K 232
 Db 151 CYSKTQKLSKIETLRLAKNYIWLSEILRSKRDPDLVSIVQTLCKGLSQPTNLVAGCLQ 210
 QY 233 SDHHHLRTAAAYEG-----GAGNATA-----AGAQ-QASGG-----SORPTPPG 270
 Db 211 LNSRNFLEQADGAGRHGSGGPPAMHPYVPCSRLAGAACQAGGLGGGAHAALRTHG 270
 QY 271 SCRTFRFAPASAGGYSVOLDALHFTSFEDSALTAMMAQKNLSPSLPGSILQPVQENSKT 330
 Db 271 YCAAYETLYAAAGGGAGSPD--YNSSEYEGPLSPPLCLNGNFS-----LKQDSSPD 319
 QY 331 SPRSHRSDGEFSPSHYS 348
 Db 320 HEKSY-----HYS 327

RESULT 12

TL5764

hypotheical protein C34E10.7 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: TL5764

R:Kirsten, J.

submitted to the EMBL Data Library, June 1994

A:Description: The sequence of C. elegans cosmid C34E10.

A:Accession number: Z18399

A:Accession: TL5764

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-192 <KIR>

A:Cross-references: EMBL:U10402; NID:g500723; PID:g500730; PIDN:AAA19069.1; CESP:C34E10.

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:C34E10.7

A:Introns: 65/2; 107/3

Query Match 10.7%; Score 200; DB 2; Length 192;

Best Local Similarity 33.8%; Pred. No. 8.5e-07;

Matches 68; Conservative 13; Mismatches 76; Indels 44; Gaps 6;

QY 157 OKORLIANARERRMHGHNHAFDQLRNVIPSFNNDKLSKYETLQMAQIYNALSELLQ 216

Db 17 RKVRVKANGREARMHGLNLDMLREYIPITTOHOKLSKIETLRLAARNYDALQRLQ 76

QY 217 TPSGGEOPPPSPACSKSDHHHLRTAAAYEGGAGNATAAGAAQASGGSORPTPGSCRTFR 276

Db 77 T---NEQPTP-----LEYAHTLANGLSQTTNNLANLLQVQRLPLPPSQFOIF 122

QY 277 SAPASAGGYSVOLDALH----FTFEDSALTAMMAQKNLSPSLPGSILQPVQENSKTSR 333

Db 123 SDPSHH-----QLHPSHPHPHSSF-----SSSPSSSCSPPPYYGPTQPS 163

QY 334 SHRSRSDGEFSP-----HSH 346

Db 164 AAPLOGSCDPQYQMYVQHSH 184

RESULT 13

JC4647

KWS protein - rat

N:Alternate names: basic helix-loop-helix protein homolog

C:Species: Rattus norvegicus (Norway rat)

C:Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 20-Jun-2000

C:Accession: JC4647

R:Kume, H.; Maruyama, K.; Tomita, T.; Iwatsubo, T.; Saido, T.C.; Obata, K.

Biochem. Biophys. Res. Commun. 219, 526-530, 1996

A:Title: Molecular cloning of a novel basic helix-loop-helix protein from the rat brain.
 A:Reference number: JC4647; MUID:96193685; PMID:8605021

A:Accession: JC4647

A:Molecule type: DNA

A:Residues: 1-381 <KUM>

A:Cross-references: DDBJ:D82868; NID:g1166397; PIDN:BAAL1615.1; PID:g1166398

A:Experimental source: brain

C:Comment: This protein is involved in synaptic plasticity, and has a role specific to n-loop-helix domain.

C:Keywords: brain

F:122-134/Region: basic

Query Match 10.7%; Score 199; DB 2; Length 381;

Best Local Similarity 26.1%; Pred. No. 2.2e-06;

Matches 81; Conservative 31; Mismatches 86; Indels 112; Gaps 12;

QY 18 DHHRQPOPHHLPQPPPP-----OPPATLOAREHPVVPPELSLLDSTDPRAWLAP 67

Db 27 DEPRSDKGADAPPQPPPPAPGSGAPPARATKCVSLRGEVP-----EP 68

QY 68 TLQIGICTARAAQYLLHSPELGASAAAPRDVDGREGELVRRSSGGASSSKSPGVKVRQ 127

Db 69 TLAEV-----KEGELGEEEEEE-EGUDE-----AEGRPKKRGPKKRWMT 112

QY 128 LCKLKGVVVDDELGCSQRAPSSQVNGVQKORLIANARERRMHGHNHAFDQLRNVIP 187

Db 113 KARLER-----SKLRQKANARERNRMDLNAALDNLKRVVP 149

QY 188 SFNNDKLSKYETLQMAQIYNALSELLQTSFGGEOPPPPPASCKSDHHHLRTAAAYEGG 247

Db 150 CYSKTQKLSKIETLRLAKNYIWLSEILRS---GKRP-----DLVSY--- 188

QY 248 AGNATAAGAAQASGGSORPTPGSCRTFRFASAGGYSVOLDALHFTFED-----SALT 303

Db 189 -----VQTLCKGLSQPT-----TNLVAGC-----LQNSRNFLEQADGAAFT 228

QY 304 AMMAQKNLSP 313

Db 229 ARVARSAMHP 238

RESULT 14

S20085

transcription factor SCL - chicken

C:Species: Gallus gallus (chicken)

C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 24-Sep-1999

C:Accession: S20085

R:Goodwin, G.; MacGregor, A.; Zhu, J.; Crompton, M.R.

Nucleic Acids Res. 20, 368, 1992

A:Title: Molecular cloning of the chicken SCL cDNA.

A:Reference number: S20085; MUID:92158629; PMID:1741264

A:Accession: S20085

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-311 <GOO>

A:Cross-references: EMBL:X63371; NID:g62844; PIDN:CAA44971.1; PID:g62845

C:Superfamily: lyl-1 protein

C:Keywords: DNA binding; transcription regulation

Query Match 10.5%; Score 195; DB 2; Length 311;

Best Local Similarity 24.6%; Pred. No. 3.2e-06;

Matches 77; Conservative 37; Mismatches 121; Indels 76; Gaps 11;

QY 29 POPPPPPQPPATLOAREHPVVPPELSLLDSTD-PRAWLAPTLOGICTARAAQYLLHSP 87

Db 7 PAPPFPSSDPR--DARRH---DPEADATSEPDSSRGGMEPPAEQOLLNGAAKEAGRPSP 61

QY 88 GASEAAAPEVDGGRG---ELVRRSSGGASSSKSPG--PVKVRQQLCKLK----- 132

Db 62 GPAAAVPVIEILVRGGSLDIKSREAAGEAMORPAGAEPCRAAEACEARVQLSPALP 121

QY 133 ----GGVVVDEILG-----CSRQAPSSKQVNG---VQKQR 161

Db 122 LPPGRAMLNVLGQPLTIGSGFFGEPSFMSYGNRVKRRPSPYEMEITDGPHTKVVR 181
QY 162 LAANARERRRMHGLNHFADOLRNVPISFNNDKKLSKYETLQMAQIYINALSELL--QTPS 219
Db 182 IFNSRRERWQNVNGFAELRKLIPTHPDDKKLSKNEILFLWKYINFLAKLNDQEE 241
QY 220 GGEQPPPPPPASCKSDHHHLRTAASYEGGAGNATAAGAAQASGSGSQRPPTPGSCRTRESAP 279
Db 242 GNQR-----GKVNKD-SGIVQEDLLQDMLSPNSSCGSSLDGA 277
QY 280 ASAGGYSVQLDAL 292
Db 278 ASDPSFTEHDTL 290

RESULT 15

JC6087
helix-loop-helix transcription factor, BETA3 - hamster
C:Species: Cricetinae Gen. sp. (hamster)
C>Date: 28-May-1997 #sequence_revision 18-Jul-1997 #text_change 05-Nov-1999
C:Accession: JC6087
R; Peyton, M.; Stellrecht, C.M.M.; Naya, F.J.; Huang, H.P.; Samora, P.J.; Tsai, M.J.
Mol. Cell. Biol. 16, 626-633, 1996
A:Title: BETA3, a novel helix-loop-helix protein, can act as a negative regulator of BETA3
A:Reference number: JC6087; MUID:96140430; PMID:8552091
A:Accession: JC6087
A:Molecule type: mRNA
A:Residues: 1-367 <PEY>
A:Cross-references: GB:S80870; NID:g1911496; PIDN:AAB50691.1; PID:g1911497
A:Experimental source: insulin tumor cell
C:Comment: This factor belongs to the tissue-specific class B basic helix-loop-helix family.
C:Genetics:
A:Gene: beta3
C:Keywords: repressor; transcription factor

Query Match 10.5%; Score 195; DB 2; Length 367;
Best Local Similarity 32.8%; Pred. No. 3.9e-06;
Matches 79; Conservative 18; Mismatches 84; Indels 60; Gaps 10;

QY 67 PTLOGICTARAAQVLLHSPBLGASEA-----APDEVVDGRGELVRRSSGGASSSKSP 119
Db 116 FNLSSL-PAGAAALCKYGESAGRGVABSSGGEGSQSPDDSDGRCELVR-AGGADPRASP 173
QY 120 GPVKVREQLCKLXGVVVDLGGCSQR-----APSSKQVN 154
Db 174 GAG-----GGGTVVE-GCSNAHLHGAGLPGSGTSGGGGGGGGGSSSKKS 223
QY 155 GVQKQRLAANARERRRMHGLNHFADOLRNVPISFNND--KKLSKYETLQMAQIYI----- 208
Db 224 KEQKALRLNINARERRRMHDLNDALDELRAVIPYAHSPSVRLSKTIALLAKNYILMQA 283
QY 209 NALSELL-----QTPSGGQPPPPPPASCKSDHHHLRTAASYEGGAGNATAAGAAQQA 259
Db 284 QALEERRLVAYLNQQAISASLPSSAAAAAALHPALGA-YEQAGYPFSAGLP 342
QY 260 S 260
Db 343 A 343

Search completed: September 21, 2004, 21:25:00
Job time : 50.2 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 20:35:26 ; Search time 27.376 Seconds
(without alignments)
673.321 Million cell updates/sec

Title: US-09-980-381A-58
Perfect score: 1862
Sequence: 1 MSRLHAEWEAEVKELGDHH.....HRSDGERSPHSHYSDSEAS 354

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1862	100.0	354	1 ATH1_HUMAN	Q92858 homo sapien
2	1627.5	87.4	351	1 ATH1_MOUSE	P48985 mus musculus
3	228.5	12.3	198	1 AMOS_DROME	Q9Y0A7 drosophila
4	222	11.9	250	1 NGN2_HUMAN	Q9H2A3 homo sapien
5	219	11.8	208	1 NGN1_BRARE	O42506 brachydanio
6	217	11.7	263	1 NGN2_MOUSE	P70447 mus musculus
7	214	11.5	352	1 NDF1_XENLA	Q91616 xenopus lae
8	213	11.4	312	1 ATO_DROME	P48987 drosophila
9	210	11.3	382	1 NDF2_HUMAN	Q15784 homo sapien
10	209.5	11.3	237	1 NGN1_HUMAN	Q92886 homo sapien
11	207.5	11.1	355	1 NDF1_MESAU	O60430 mesocricetu
12	205.5	11.0	337	1 NDF6_MOUSE	P48986 mus musculus
13	205	11.0	244	1 NGN1_RAT	P70595 rattus norv
14	204	11.0	382	1 NDF2_RAT	Q63689 rattus norv
15	202	10.8	244	1 NGN1_MOUSE	P70660 mus musculus
16	202	10.8	357	1 NDF1_MOUSE	Q60867 mus musculus
17	202	10.8	357	1 NDF1_RAT	Q64289 rattus norv
18	202	10.8	383	1 NDF2_MOUSE	P62414 mus musculus
19	201.5	10.8	214	1 NGN3_MOUSE	P70661 mus musculus
20	200.5	10.8	337	1 NDF6_HUMAN	Q96N88 homo sapien
21	200	10.7	192	1 YLB7_CAEEL	P46581 caenorhabdi
22	199	10.7	356	1 NDF1_HUMAN	Q13562 homo sapien
23	197.5	10.6	357	1 NDF1_CHICK	P79765 gallus gall
24	195	10.5	311	1 SCL_CHICK	P24899 gallus gall
25	195	10.5	367	1 BET3_MESAU	O09029 mesocricetu
26	192.5	10.3	214	1 NGN3_HUMAN	Q9Y422 homo sapien
27	191.5	10.3	330	1 NDFM_CHICK	P79766 gallus gall
28	191.5	10.3	331	1 NDF4_HUMAN	Q9hd90 homo sapien
29	187	10.0	330	1 NDF4_MOUSE	O09105 mus musculus
30	184	9.9	71	1 L132_CAEEL	Q10574 caenorhabdi
31	183	9.8	298	1 OLG2_CHICK	Q902b3 gallus gall
32	179	9.6	331	1 TAL_HUMAN	P17542 homo sapien
33	178	9.6	329	1 TAL_MOUSE	P22091 mus musculus

RESULT 1

ATH1_HUMAN

ID ATH1_HUMAN STANDARD; PRT; 354 AA.

AC Q92858;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Atonal protein homolog 1 (Helix-loop-helix protein hATH-1).

GN ATOH1 OR ATH1.

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

[1]

RP SEQUENCE FROM N.A.

RX MEDLINE=97026280; PubMed=8872459;

RA Ben-Arie N., McCall A.E., Berkman S., Eichele G., Bellien H.J.,

RA Zoghbi H.Y.;

RT "Evolutionary conservation of sequence and expression of the bHLH

protein Atonal suggests a conserved role in neurogenesis.";

RL Hum. Mol. Genet. 5:1207-1216(1996).

CC -!- FUNCTION: Activates E box-dependent transcription in collaboration

with E47, but the activity is completely antagonized by the

negative regulator of neurogenesis HES1. May play a role in the

differentiation of subsets of neural cells by activating E box-

dependent transcription (By similarity).

CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another

bHLH protein.

CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).

CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: U61148; AAB41305.1; -

CC TRANSFAC: T04544; -

CC Genew; HGNC:797; ATOH1.

CC MIM; 601461; -

CC GO; GO:0003700; P:transcription factor activity; TAS.

CC GO; GO:0007417; P:central nervous system development; TAS.

CC GO; GO:0006366; P:transcription from Pol II promoter; TAS.

CC InterPro; IPR001092; HLH_basic.

CC Pfam; PF00010; HLH; 1.

CC SMART; SM00353; HLH; 1.

CC PROSITE; PS50988; HLH; 1.

CC Transcription regulation; Activator; DNA-binding; Nuclear protein.

FT DOMAIN 29 38 POLY-PRO.

FT DNA_BIND 160 171 BASIC DOMAIN.

FT DOMAIN 172 212 HELIX-LOOP-HELIX MOTIF.

FT DOMAIN 224 228 POLY-PRO.

FT SEQUENCE 354 AA; 38160 MW; AB12F1E917A00A8D CRC64;

P12980 homo sapien
P79920 xenopus lae
P27792 mus musculus
Q13516 homo sapien
O35885 mus musculus
O16867 drosophila
Q9eqw6 mus musculus
Q64124 mus musculus
Q12870 homo sapien
Q60539 mesocricetu
Q60756 mus musculus
O60682 homo sapien

ALIGNMENTS

```
Query Match 100.0%; Score 1862; DB 1; Length 354;
Best Local Similarity 100.0%; Pred. No. 5.9e-110;
Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSRLHAEWAWEVKELGDHHRQPPHLPQPPPPQPPATLOAREHVPVPELSDSTD 60
Db 1 MSRLHAEWAWEVKELGDHHRQPPHLPQPPPPQPPATLOAREHVPVPELSDSTD 60

QY 61 PRAWLPTLOGICTARAAQVLLHSPELGASEAAAPRDEVGRGELVRRSSGGASSKSPG 120
Db 61 PRAWLPTLOGICTARAAQVLLHSPELGASEAAAPRDEVGRGELVRRSSGGASSKSPG 120

QY 121 PVKVRQOLCKLKGVVVDELGCQRAPSSKQVNGVQKRRLAANARERRMGLNHPD 180
Db 121 PVKVRQOLCKLKGVVVDELGCQRAPSSKQVNGVQKRRLAANARERRMGLNHPD 180

QY 181 QLRNVIPSNNDKLSKYETLQMAQIYINALSELLQTPSGGEGPPPPASCKSDHHHLRT 240
Db 181 QLRNVIPSNNDKLSKYETLQMAQIYINALSELLQTPSGGEGPPPPASCKSDHHHLRT 240

QY 241 AASVEGAGNATAAGAAQASGSGSORPPGSCRTFRSAPASAGYSVQLDHFSTFEDS 300
Db 241 AASVEGAGNATAAGAAQASGSGSORPPGSCRTFRSAPASAGYSVQLDHFSTFEDS 300

QY 301 ALTAMMAQKNLSPSLPGSILQPVQENSKTSPRSHRSDGFEFSPHSHYSDDSEAS 354
Db 301 ALTAMMAQKNLSPSLPGSILQPVQENSKTSPRSHRSDGFEFSPHSHYSDDSEAS 354

RESULT 2
ATH1 MOUSE STANDARD; PRT; 351 AA.
AC P48985;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Atonal protein homolog 1 (Helix-loop-helix protein MATH-1) (MATH1).
GN ATOH1 OR ATH1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=129/J;
RC MEDLINE=95238366; PubMed=7721778;
RX Akazawa C., Ishibashi M., Shimizu C., Nakanishi S., Kageyama R.;
RT "A mammalian helix-loop-helix factor structurally related to the
RT product of Drosophila proneural gene atonal is a positive
RT transcriptional regulator expressed in the developing nervous
RT system.";
RL J. Biol. Chem. 270:8730-8738 (1995).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932;
RX Strausberg R.L., Reingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins S.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Tohyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
```

```
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -1- FUNCTION: Activates E box-dependent transcription in collaboration
CC with E47, but the activity is completely antagonized by the
CC negative regulator of neurogenesis HES1. May play a role in the
CC differentiation of subsets of neural cells by activating E box-
CC dependent transcription.
CC -1- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein.
CC -1- TISSUE SPECIFICITY: Developing nervous system, and in adult
CC epithelial cells of the gastrointestinal tract.
CC -1- DEVELOPMENTAL STAGE: First detected in the cranial ganglions and
CC the dorsal part of the central nervous system on embryonic day 9.5
CC (E9.5). From E10.5 onward, prominent expression of MATH-1
CC continues in the dorsal part of the central nervous system but
CC becomes restricted to the external granular layer of the
CC cerebellum by e18 and is undetectable in the adult nervous system.
CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D43694; BAA07791.1; -.
CC DR EMBL; BC010820; AAH10820.1; -.
CC DR EMBL; BC051256; AAH51256.1; -.
CC DR PIR; A56387; A56387.
CC DR TRANSFAC; T01668; -.
CC DR MGD; MGI:104654; Atohl.
CC DR GO; GO:0007420; P:brain development; IMP.
CC DR InterPro; IPR001092; HLH_basic.
CC DR Pfam; PF00010; HLH; 1.
CC DR SMART; SM00353; HLH; 1.
CC DR PROSITE; PS00888; HLH; 1.
CC KW Transcription regulation; Activator; DNA-binding; Nuclear protein.
CC FT DNA_BIND 157 168 BASIC DOMAIN.
CC FT DOMAIN 169 209 HELIX-LOOP-HELIX MOTIF.
CC FT DOMAIN 221 224 POLY-PRO.
CC SQ SEQUENCE 351 AA; 37854 MW; 596503DF23C3BF96 CRC64;
Query Match 87.4%; Score 1627.5; DB 1; Length 351;
Best Local Similarity 88.4%; Pred. No. 2.9e-95;
Matches 313; Conservative 11; Mismatches 27; Indels 3; Gaps 2;

QY 1 MSRLHAEWAWEVKELGDHHRQPPHLPQPPPPQPPATLOAREHVPVPELSDSTD 60
Db 1 MSRLHAEWAWEVKELGDHHRQPPHLPQPPPPQPPATLOAREHVPVPELSDSTD 58

QY 61 PRAWLPTLOGICTARAAQVLLHSPELGASEAAAPRDEVGRGELVRRSSGGASSKSPG 120
Db 59 PRAWLPTLOGICTARAAQVLLHSPELGASEAAAPRDEADSQGLVYR-SGCGLSKSPG 117

QY 121 PVKVRQOLCKLKGVVVDELGCQRAPSSKQVNGVQKRRLAANARERRMGLNHPD 180
Db 118 PVKVRQOLCKLKGVVVDELGCQRAPSSKQVNGVQKRRLAANARERRMGLNHPD 177

QY 181 QLRNVIPSNNDKLSKYETLQMAQIYINALSELLQTPSGGEGPPPPASCKSDHHHLRT 240
Db 178 QLRNVIPSNNDKLSKYETLQMAQIYINALSELLQTPNVGEGPPPTASCKNDHHHLRT 237

QY 241 AASVEGAGNATAAGAAQASGSGSORPPGSCRTFRSAPASAGYSVQLDHFSTFEDS 300
Db 238 ASSVEGAGASAVAGAPAGGGPRPTPGPCRTFRSGSGSGYSVQLDHFPAFEDR 297

QY 301 ALTAMMAQKNLSPSLPGSILQPVQENSKTSPRSHRSDGFEFSPHSHYSDDSEAS 354
```

Db 298 ALTAMMAKOLSPSLPGILQVQEDNSKTSRSHRSDGEPSPHSHYSDSEAS 351

RESULT 3

ID AMOS DROME STANDARD; PRT; 198 AA.

AC Q9Y0A7; Q9VJ76;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Basic helix-loop-helix transcription factor Amos (Reduced olfactory

DE organs protein) (Rough eye protein) (Absent MD neurons and olfactory

DE sensilla protein) (Amos protein).

GN AMOS OR ROI OR ROLO OR CG10393.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A., AND FUNCTION.

RC STRAIN=Oregon-R;

RC MEDLINE=20170246; PubMed=10707973;

RA "Amos, a proneural gene for Drosophila olfactory sense organs that is

RT regulated by lozenge."

RL Neuron 25:69-78(2000).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Berkley;

RC MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berhan P.P., Bhandari D., Bolshakov S.,

RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Fostler K.E., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster."

RL Science 287:2185-2195(2000).

RN [3]

RP FUNCTION.

RC MEDLINE=20170245; PubMed=10707972;

RA Huang M.L., Hsu C.H., Chien C.T.;

RT

"The proneural gene amos promotes multiple dendritic neuron formation in the Drosophila peripheral nervous system.";

Neuron 25:57-67(2000).

-!- FUNCTION: Transcription factor involved in early neurogenesis.

CC Promotes multiple dendritic (MD) neuron formation. Required for

CC olfactory sensilla.

CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another

CC bHLH protein. Interacts with Daughterless.

CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).

CC -!- DEVELOPMENTAL STAGE: During embryonic development, amos is

CC expressed in patches of ectodermal cells, and the expression is

CC quickly restricted to sensory organ precursors.

CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or

CC send an email to license@isb-sib.ch).

CC

EMBL; AF166113; AAD45410.1; -;

DR EMBL; AE003659; AAF53678.1; -;

DR HSSP; P25912; 1HLO.

DR FlyBase; FBgn0003270; amos.

DR InterPro; IPR001092; HLH_basic.

DR Pfam; PF00010; HLH; 1.

DR SMART; SM00353; HLH; 1.

DR PROSITE; PS00888; HLH; 1.

DR Nuclear protein; Transcription regulation; Developmental protein;

KW Neurogenesis.

FT DOMAIN 111 128 SER-RICH

FT DNA_BIND 138 150 BASIC DOMAIN

FT DOMAIN 151 191 HELIX-LOOP-HELIX MOTIF.

FT CONFLICT 25 25 S -> G (IN REF. 1).

SQ SEQUENCE 198 AA; 22569 MW; FID22AF2BA1C670B CRC64;

Query Match 12.3%; Score 228.5; DB 1; Length 198;

Best Local Similarity 34.6%; Pred. No. 6.5e-08;

Matches 73; Conservative 22; Mismatches 57; Indels 59; Gaps 7;

QY 34 PPQPPA-----TLQAREHPVPPPELSLDSTDPRA-----WLAPT---LOGICT 74

Db 14 PDEAPAEFLNSDFFQLEQLMYQEFSTDSQSGANGSCSLEMYDYTPSVLEHMLN 73

QY 75 AR-AAQYLLHSPFELGASAAAPR-----DEVDCRGELVRRSSGGASSKSPGVKVR 125

Db 74 AQEQOQHHLQANPLGNQGRSPRYWNKQQRKPYDKLSTSMSSSTSSASSSS----- 126

QY 126 EQLCKLKGVVVDLGCGRAPSSKQVNG-VOKORRLAANARERRRMHGLNHPDQLRN 184

Db 127 -----SSAGTGGVVKGRRLAANARERRRMHGLNHPDQLRN 163

QY 185 VIPSPFNNDKLSKYETLQMAQIYINALSELL 215

Db 164 VVPSLGHDRRLSKYETLQMAQYIGDLVTL 194

RESULT 4

NGN2 HUMAN STANDARD; PRT; 250 AA.

ID NGN2 HUMAN

AC Q9H2A3;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Neurogenin 2 (Fragment).

GN NEUROG2 OR NGN2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

```
RP SEQUENCE FROM N.A.
RA Simmons A.D., Horton S., Abney A.L., Johnson J.E.;
RT "Neurogenin 2 expression in ventral and dorsal spinal neural tube
RL progenitor cells is regulated by distinct enhancers.";
RT Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Involved in neurogenesis (By similarity).
CC -1- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF303002; AAG40770.1; -.
CC Genew; HGNC:13805; NEUROG2.
CC MIM; 606624; -.
CC InterPro; IPR001092; HLH_basic.
CC Pfam; PF00010; HLH; 1.
CC SMART; SM00353; HLH; 1.
CC PROSITE; PS50888; HLH; 1.
CC Neurogenesis; Developmental protein, Differentiation.
CC NON_TER 1
CC DNA_BIND 100 111 BASIC DOMAIN.
CC DOMAIN 112 152 HELIX-LOOP-HELIX MOTIF.
CC SEQUENCE 250 AA; 26130 MW; EB381FF5EF1500EF CRC64;

Query Match 11.9%; Score 222; DB 1; Length 250;
Best Local Similarity 30.4%; Pred. No. 2.1e-07;
Matches 90; Conservative 27; Mismatches 89; Indels 90; Gaps 14;

QY 84 SPELGA-----SEAAAPRDVDRGELVLRSSGGASSKSPGVKVRQELCKLG--GVV 136
DB 2 SPALAAITPLSSADEEEEPGASGARRQRAAGAGGVAAGACRPARLLGLV 61
QY 137 VBLEGC-----SKQNGVQKORRLAANAFERRRMHGLHAFDQLRNVIP 187
DB 62 HD---CKRPPSRAVRSRGAETVQRIKTKRLKANNRRNRNMENLNAALDREVLP 118
QY 188 SFNNDKLSKYETLQMAQIYNALSELLQTPSGEGQPPPPASCKSDHHHLRTAAVEGG 247
DB 119 TFPEDAKTKIETLRFAHYIINALTETLR-----LADHCGG 154
QY 248 AGNAT-----AAGAQAQAGSGQRPTPPGSCRTFRSAPAGAGGYVQLDAL 292
DB 155 GGGGLPGALFSEAVLLSPGASAAALSSGDSFSPASTWSC-TNSPAPSS----- 203
QY 293 HSTFSDSALTAMAAQKNLSPSL-GSIL---QPVEENSKYSPRHSRSDGRFSPH 344
DB 204 -VSSNSTSPSYSC-----TLSPASPGSDMDYQWQ-----PPDKHR-----YAPH 242

RESULT 5
NGN1_BRARE
ID NGN1_BRARE STANDARD; PRT; 208 AA.
AC O42606;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neurogenin 1 (Neurogenic differentiation factor 3) (NeuroD3)
DE (Neurogenin related protein-1).
GN NEUROG1 OR NEUROD3 OR NGN1 OR NGRI.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
```

```
[1]
RP SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, AND DEVELOPMENTAL
RP STAGE.
RX MEDLINE=98072331; PubMed=9409673;
RA Blader P., Fischer N., Gradwohl G., Guillemont F., Straehle U.;
RT "The activity of neurogenin1 is controlled by local cues in the
RT zebrafish embryo.";
RL Development 124:4557-4569 (1997).
[2]
RP SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, AND DEVELOPMENTAL
RP STAGE.
RX TISSUE=Embryo;
MEDLINE=98129571; PubMed=9469669;
RA Kim C.-H., Bae Y.-K., Yamanaka Y., Yamashita S., Shimizu T., Fujii R.,
RA Park H.-C., Yeo S.-Y., Huh T.-L., Hibi M., Hirano T.;
RT "Overexpression of neurogenin induces ectopic expression of HuC in
RT zebrafish.";
RL Neurosci. Lett. 239:113-116 (1997).
[3]
RP SEQUENCE FROM N.A.
RA Thisse B.;
RT "Danio rerio neurogenin the earliest marker of primary neurons starts
RT to be expressed at midgastrula.";
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
RX TISSUE=Embryo;
MEDLINE=98401940; PubMed=9733104;
RA Korzh V., Sleptsova I., Liao J., He J., Gong Z.;
RT "Expression of zebrafish bHLH genes ngn1 and nrd defines distinct
RT stages of neural differentiation.";
RL Dev. Dyn. 213:92-104 (1998).
CC -1- FUNCTION: Appears to mediate neuronal differentiation.
CC -1- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein (By similarity).
CC -1- TISSUE SPECIFICITY: Embryonic nervous system.
CC -1- DEVELOPMENTAL STAGE: First detected at the late gastrulation stage
CC (9 hours). Strongly expressed in distinct domains in the neural
CC place at the 3-somite stage. At 24 hours, expressed in specific
CC regions of the developing brain and in the spinal cord.
CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF024535; AAB82271.1; -.
CC EMBL; AF017301; AAB70528.1; -.
CC EMBL; AF036149; AAB88821.1; -.
CC EMBL; U94588; AAD09254.1; -.
CC ZFIN; ZDB-GENE-990415-174; neurogl.
CC InterPro; IPR001092; HLH_basic.
CC Pfam; PF00010; HLH; 1.
CC SMART; SM00353; HLH; 1.
CC PROSITE; PS50888; HLH; 1.
CC DNA-binding; Nuclear protein; Transcription regulation; Activator;
CC Neurogenesis; Developmental protein; Differentiation.
CC DNA_BIND 71 82 BASIC DOMAIN.
CC DOMAIN 83 123 HELIX-LOOP-HELIX MOTIF.
CC SEQUENCE 208 AA; 22911 MW; CDB333AE483DC51A CRC64;

Query Match 11.8%; Score 219; DB 1; Length 208;
Best Local Similarity 30.1%; Pred. No. 2.7e-07;
Matches 69; Conservative 32; Mismatches 76; Indels 52; Gaps 8;

QY 97 DEVDGRGELVRRSSGGASSKSPGVKVRQELCKLGKGVGVVDELGCSQRAPSSKQNGV 156
DB 21 DDEDSRSSL-HPASPASSCGKPPASGLQKKRRRG-----RARNETTTHVV 67
```

QY 157 QKORLAANARRRRMHGHNHAFDOLRNVPSPFNNDKKLSKYETLQMAQIYINALSELLO 216
 DB 68 KKNRRLKANDRRNRHNLNDALDRLSVLPAPDDTKLTIKTETURFAHNYIWAUSETIR 127
 QY 217 ---TPSGGGEQPPPP---PPASCKSDHHLRLTAASYEGGAGNATAAGAAQA---SGSGSRPT 267
 DB 128 IADQKQKSRDCPLLLPLGLSCWAD-----APSPGSDSCSWPFGASSSSS 171
 QY 268 PPGSCRTFRSPASAGGYS-VOLDALHFTFEDSALTAMMAQKNLSPLS 315
 DB 172 SPSCNSDPGSPAAAMDFFGLQTDVY-----SYRNFVPSI 207

RESULT 6
 NGN2 MOUSE
 ID NGN2 MOUSE STANDARD; PRT; 263 AA.
 AC P70447; P70237;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Neurogenin 2 (Atonal protein homolog 4) (Helix-loop-helix protein
 DE MATH-4A) (MATH4A).
 GN NEUROG2 OR NGN2 OR ATOH4 OR ATH4A.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97153565; PubMed=9000438;
 RA Sommer L., Ma Q., Anderson D.J.,
 RT "Neurogenins, a novel family of atonal-related bHLH transcription
 RT factors, are putative mammalian neuronal determination genes that
 RT reveal progenitor cell heterogeneity in the developing CNS and PNS."
 RL Mol. Cell. Neurosci. 8:221-241(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CD-1;
 RX MEDLINE=971105839; PubMed=8948587;
 RA Gradwohl G., Fode C., Guillemot F.,
 RT "Restricted expression of a novel murine atonal-related bHLH protein
 RT in undifferentiated neural precursors."
 RL Dev. Biol. 180:227-241(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129;
 RA Simmons A.D., Horton S., Abney A.L., Johnson J.E.,
 RT "Neurogenin 2 expression in ventral and dorsal spinal neural tube
 RT progenitor cells is regulated by distinct enhancers."
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Involved in neurogenesis.
 CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
 CC bHLH protein.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U76207; AAC53028.1; -
 CC EMBL; Y07621; CAAG8900.1; -
 CC EMBL; AF303001; AAC40769.1; -
 CC MGD; MGI:109619; Neurog2.
 CC InterPro; IPR001092; HLH_basic.
 CC Pfam; PF00010; HLH; 1.
 CC SMART; SM00353; HLH; 1.
 CC PROSITE; P55088; HLH; 1.

KW DNA-binding; Nuclear protein; Transcription regulation; Activator;
 KW Neurogenesis; Developmental protein; Differentiation.
 FT DNA_BIND 113 124 BASIC DOMAIN.
 FT DOMAIN 125 165 HELIX-LOOP-HELIX MOTIF.
 FT CONFLICT 60 E -> G (IN REF. 2).
 SQ SEQUENCE 263 AA; 28215 MW; 817EF8246BD8CABE CRC64;

Query Match 11.7%; Score 217; DB 1; Length 263;
 Best Local Similarity 28.7%; Pred. No. 4.6e-07;
 Matches 98; Conservative 25; Mismatches 89; Indels 130; Gaps 15;

QY 40 TLQAREHFVYPPELSLLDSTDTP-RAWLAP-----TLQGITAPAAQYL 81
 DB 7 TLELKEE---BEVLMILGASPASATLTPMSSADEEDEELRRPGSARGQGAEEQGV 63
 QY 82 LHSPELGASEAAAPRDEVDGELVRRSGGASSSKSPGVKVRQLCKLKGVVVDELG 141
 DB 64 QGSPASGA-----GGC-----REFRLGLMHECKRRP----- 90
 QY 142 CSRQRAPS-----SKQVNGVQKORBLAANARRRRMHGHNHAFDOLRNVPFNNDKKLS 196
 DB 91 -SRSAVRGAKTAETVQRIKTKRKKANNRRNRHNLNAALDALREVLPFPEDAKLT 149
 QY 197 KYETLQMAQIYINALSELLOTPSGGGEQPPPPPPASCKSDHHLRLTAASYEGGAGNATAA-- 254
 DB 150 KIETLRFHNYIWAUSETIR-----LADHCAGAGGLQALF 185
 QY 255 -----GAQOASGGGQRPTTPGSCRTFRSPASAGYSVOLDALHFTFEDSALTAM 305
 DB 186 TEAVLLSPGAALGAGSDSPSPSSWSCT---NSPASSN-----STSPYS- 228
 QY 306 MAQKNLSPLSGSIL---QPVOEENSKTSPRSHRSDGFSPH 344
 DB 229 ----TLSPASGSDVDYWP-----PPPEKHR-----YAPH 255

RESULT 7
 NDPL_XENLA
 ID NDPL_XENLA STANDARD; PRT; 352 AA.
 AC Q91616;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Neurogenic differentiation factor 1 (NeuroD1).
 GN NEUROD1 OR NEUROD.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95273957; PubMed=7754368;
 RA Lee J.E., Hollenberg S.M., Snider L., Turner D.L., Lipnick N.,
 RA Weintraub H.,
 RT "Conversion of Xenopus ectoderm into neurons by NeuroD, a basic
 RT helix-loop-helix protein."
 RL Science 268:836-844(1995).
 CC -!- FUNCTION: Acts as a differentiation factor during neurogenesis.
 CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
 CC bHLH protein.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -!- TISSUE SPECIFICITY: Expressed in differentiating neurons of
 CC both the central and peripheral nervous systems.
 CC -!- DEVELOPMENTAL STAGE: Expressed during embryonic development.
 CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

```

CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U28067; AAC59675.1; -.
DR PIR; I51687; I51687.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS50888; HLH; 1.
DR DNA-binding; Nuclear protein; Transcription regulation; Activator;
KW Neurogenesis; Developmental protein; Differentiation.
FT DOMAIN 59 78 GLU-RICH (ACIDIC).
FT DOMAIN 88 94 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 103 114 BASIC DOMAIN.
FT DOMAIN 115 155 HELIX-LOOP-HELIX MOTIF.
SQ SEQUENCE 352 AA; 39662 MW; 226298DB348233E CRC64;

Query Match 11.58; Score 214; DB 1; Length 352;
Best Local Similarity 29.5%; Pred. No. 9.9e-07;
Matches 70; Conservations 31; Mismatches 86; Indels 50; Gaps 9;

QY 144 RORAPSKQVNGVQ---KORLAANARERRRHGLHAFDQLNRNVPSPNNKKLSKYE 199
DB 83 KRGPKKKKTKARVERFKYRMKANARERNHGLNDALDSLKVKVPCYSKTKQLSKIE 142
QY 200 TLOWAIYINALSELL-----QTPSGEQPPPP-----PASCK 232
DB 143 TLRLAKNYIWNALSEILRSKSPDLVSFVQTLKGLSQPTTNLVAGCLQLNPRTPFLPQSQ 202
QY 233 SDHHHLRTAASYEGGAGNATAGAAQAGSGSQPTTPP-GSCRTFFSPAPAGGYVQLDA 291
DB 203 DIOGHMGTASS-----SFLQGYVYQSPG--LPSPPYGTMDSHVFHVKPHSGAALP 254
QY 292 LHFSTFSDSALTAMMAQKNLSPLSGSILOPVOEENSKTSPRSHRS-DGEFFSPHSY 347
DB 255 F-----FDSSIVT-----ECTSPSFDGLPFLSLVNGNFTFKHESEYDKNYTTHY 302

RESULT 8
ATO DROME STANDARD; PRT; 312 AA.
AC P48987; Q9VHU0.
DT 01-FEB-1996 (Rel. 33, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Atonal protein.
GN ATO OR CG7508.
OS Drosophila melanogaster (Fruit fly)
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RN SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
RC STRAIN=Oregon-R;
RC MEDLINE=93313961; PubMed=8324823;
RX Jarman A.P., Grau Y., Jan L.Y., Jan Y.N.;
RA "Atonal is a proneural gene that directs chordotonal organ formation
RT in the Drosophila peripheral nervous system.";
RL Cell 73:1307-1321(1993).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=Berkelley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.P.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Stutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktoglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Borchan M.R., Bouck J., Brokstein P., Brottier P.,

```

```

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [3]
RN FUNCTION.
RP MEDLINE=94255014; PubMed=8196767;
RX Jarman A.P., Grell E.H., Ackerman L., Jan L.Y., Jan Y.N.;
RA "Atonal is the proneural gene for Drosophila photoreceptors.";
RT Nature 369:398-400(1994).
CC -!- FUNCTION: Developmental protein involved in neurogenesis. Required
CC for the formation of chordotonal organs and photoreceptors. Seems
CC to bind to E boxes. Specifically required for the photoreceptor R8
CC selection.
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein. Forms a heterodimer with Daughterless.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Proneural clusters and sense organ precursors
CC of the chordotonal organs, optic tectum of the eye-antennal disk
CC and developing brain lobe.
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L36646; AAA21879.1; -.
DR EMBL; AE003678; AAF54209.1; -.
DR PIR; A40708; A40708.
DR FlyBase; FBgn0010433; ato.
DR GO; GO:0007420; P:brain development; IMP.
DR GO; GO:0007173; P:EGF receptor signaling pathway; IGI.
DR GO; GO:0007605; P:hearing; IMP.
DR GO; GO:0007438; P:neocortical nervous system development; IMP.
DR GO; GO:0007422; P:peripheral nervous system development; NAS.
DR GO; GO:0045464; P:R8 cell fate specification; NAS.
DR GO; GO:0007224; P:smoothed signaling pathway; IGI.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS50888; HLH; 1.
KW Neurogenesis; Differentiation; Developmental protein; Nuclear protein;
KW Transcription regulation; DNA-binding
FT DNA_BIND 255 267 BASIC DOMAIN
FT DOMAIN 268 308 HELIX-LOOP-HELIX MOTIF.

```

FT CONFLICT 149 149 G -> A (IN REF. 1).
SQ SEQUENCE 312 AA; 34116 MW; 069479287438F456 CRC64;

Query Match
Best Local Similarity 11.4%; Score 213; DB 1; Length 312;
Matches 68; Conservative 21; Mismatches 85; Indels 26; Gaps 6;

QY 33 PPQPQA-----TLQAREHVPVPELSLLDSTD--RAWLAPTLOGICTARA 77
DB 123 PVASPPAVEMGSSNVGCTKTPASAGP--KPRSVYTKNQKQSTTATSTPTAAESSASV 180
QY 78 AOYLLHSPELGASEAAAPRDEVDGRLVRRSSGGASSKSPGPVKVREQLCKLKGVVV 137
DB 181 NLYTEEFQNFDFNSALFDDSDVEDDLN-LFSGGEDFDNDGSF-----DLADGENQ 232
QY 138 DELGCSQRAPSKQVNG-VQKORLAANARERRMHGLNHFADQLRVNVPFNNDKKLS 196
DB 233 DAAAGSGGKRRGKQITPVVKKRRLAANARERRRMQNLQAFDLRLQYLPCLGNDRLS 292
QY 197 KYETLQMAQIYINALSELQ 216
DB 293 KHEITLQMAQIYINALGDLRL 312

RESULT 9
ID NDF2 HUMAN STANDARD; PRT; 382 AA.
AC Q15784;
DT 01-NOV-1997 (Rel. 35, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neurogenic differentiation factor 2 (NeuroD2).
GN NEUROD2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96413331; PubMed=8816493;
RA McCormick M.B., Tamimi R.M., Snider L., Asakura A., Bergstrom D.,
RA Tapscott S.J.;
RT "NeuroD2 and neuroD3: distinct expression patterns and
RT transcriptional activation potentials within the neuroD gene
RT family.";
RL Mol. Cell. Biol. 16:5792-5800 (1996).
RN [2]
RP REVISIONS TO 336-338.
RA Tapscott S.J., Tamimi R.T., McCormick B.M.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Appears to mediate neuronal differentiation.
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U58681; AAC26058.1; -.
DR TRANSFAC; T04903; -.
DR Genew; HGNC:7763; NEUROD2.
DR MIM; 601725; -.
DR GO; GO:0003700; F:transcription factor activity; TAS.
DR GO; GO:0007399; P:neurogenesis; TAS.
DR GO; GO:0006357; P:regulation of transcription from Pol II promoter; TAS.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.

DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00888; HLH; 1.
KW Transcription regulation; DNA-binding; Nuclear protein;
KW Differentiation.
FT DOMAIN 37 46 POLY-PRO.
FT DOMAIN 82 91 POLY-GLU.
FT DOMAIN 107 113 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DNA_BIND 122 133 BASIC DOMAIN.
FT DOMAIN 134 174 HELIX-LOOP-HELIX MOTIF.
FT DOMAIN 282 285 POLY-GLY.
SQ SEQUENCE 382 AA; 41361 MW; E33AA5C0BE371B47 CRC64;

Query Match
Best Local Similarity 11.3%; Score 210; DB 1; Length 382;
Matches 98; Conservative 41; Mismatches 119; Indels 110; Gaps 17;

QY 22 QQPHLPPPPPPPPATLQAREHVPVPELSLLDSTDPRAWLAPTLOGICTARAQYL 81
DB 28 EPRSDKGADPPPPPPAP-----GPGAGP--PARAAKPV 58
QY 82 LHSPELGASEAAAPRDEVDGRLVRRSSGGASSKSPGPVKVREQLCKLKGVVVDELG 141
DB 59 ----PLRGEETATLAERKEGEL-----GGEEREE-----EKEGLDAEG-----E 99
QY 142 CSQRAPSSQVNGVQ----KORLAANARERRMHGLNHFADQLRVNVPFNNDKKLSK 197
DB 100 RPKRGPKKRMKTKARLERSKLRRKANARERRMHDLNALDNLKRVVPCYSKTKLSK 159
QY 198 YETLQMAQIYINALSELL-----QTPSGEQPPPP--PASC-KSDHHHLRTAA 242
DB 160 IETLRKNIWALSEILRSKRPDLVSVVQTLCKGLSQPTTNLVAGLQLNSRNFLTEQ 219
QY 243 SYEG-----GAGNATA-----AGAQ-OASGG-----SQRPPTPGSCETRESAPA 280
DB 220 GADGAGRFGHSGGPFAMHPYPYPCSLAGACQAAAGLGGGAHALRTGYCAAYETLVA 279
QY 281 SAGYSVQLDALHFTFEDSALTAMAAKNLSPSLPGSILQPVQENSKTSRSHRSDEG 340
DB 280 AAGGGGASPD-YNSSEYEGPLPPLCLNGNFS-----LKQDSSPDHEKSY----- 323
QY 341 FSPHSHYS 348
DB 324 ----HYS 326

RESULT 10
ID NGN1 HUMAN STANDARD; PRT; 237 AA.
AC Q92886; Q96HE1;
DT 15-DEC-1998 (Rel. 37, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neurogenin 1 (Neurogenic differentiation factor 3) (NeuroD3)
DE (Neurogenic basic-helix-loop-helix protein).
GN NEUROG1 OR NGN1 OR NGN OR NEUROD3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96413331; PubMed=8816493;
RA McCormick M.B., Tamimi R.M., Snider L., Asakura A., Bergstrom D.,
RA Tapscott S.J.;
RT "NeuroD2 and neuroD3: distinct expression patterns and
RT transcriptional activation potentials within the neuroD gene
RT family.";
RL Mol. Cell. Biol. 16:5792-5800 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSEB=Brain, and Eye;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

```

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA R. Richards S., Worley K.C., Haje S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Butterfield Y.S.N., Grimwood J., Schmutz J., Myers R.M.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A., Smailus D.E.,
RT "generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Appears to mediate neuronal differentiation.
CC -1- SUBUNIT: Efficient DNA binding requires dimerization with another
CC BHLH protein.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- TISSUE SPECIFICITY: Expression restricted to the embryonic
CC nervous system.
CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U63842; AAB37575.1; -.
CC DR EMBL; BC008687; AAH08687.1; -.
CC DR EMBL; BC028226; AAH28226.1; -.
CC DR TRAFAC; T04907; -.
CC DR Genew; HGNC:7764; NEUROG1.
CC DR MIM; 601726; -.
CC DR GO; 0003700; F:transcription factor activity; TAS.
CC DR GO; 0007399; P:neurogenesis; TAS.
CC DR GO; 0006357; P:regulation of transcription from Pol II pro. .; TAS.
CC DR InterPro; IPR001092; HLH_basic.
CC DR Pfam; PF00010; HLH; 1.
CC DR SMART; SM00353; HLH; 1.
CC DR PROSITE; PS00888; HLH; 1.
CC DR DNA-binding; Nuclear protein; Transcription regulation; Activator;
CC Neurogenesis; Developmental protein; Differentiation.
CC FT DNA BIND 93 104 BASIC DOMAIN.
CC FT DOMAIN 105 145 HELIX-LOOP-HELIX MOTIF.
CC FT CONFLICT 51 51 G->S (IN REF.1).
CC FT SEQUENCE 237 AA; 25717 MW; E3702A0B3408D567 CRC64;
CC
CC Query Match 11.3%; Score 209.5; DB 1; Length 237;
CC Best Local Similarity 33.8%; Pred. No. 1.2e-06;
CC Matches 70; Conservative 26; Mismatches 64; Indels 47; Gaps 9;
CC
CC QY 107 RRSAGAS-SKSPGVKVEQICLKGVVVDLGCGRORAPSSKQVGVQKRIIAAN 165
CC Db : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 49 REGAPNISRASVEGAQDDQERRRRG-----RTRVRSALLSLRRSRVKAN 98
CC
CC QY 166 ARRRRHGLNHPDOLRNVPISFNNDKLSKVETLQMAQIYINALSELQ-----TPSG 220
CC Db : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 99 DRENRMHNLNADLRSVLPFPDDTKLTKETLRFANVYIWAETLRLADQLPGG 158
CC
CC QY 221 GEQ-----PPASCKSDHHHLRTAASVEGGAGNATAAGAAQAGSGSQ-----RP 266
CC Db : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 159 GARERLLPPQCVCLPGPPSPASD-----AESWGSAGAAASPLSDSPSPAAASEDTYRP 212
CC
CC QY 267 TPQSCRTSRAPASAGYSVQDLALH 293
CC Db : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db 213 GDP-----VESFP-----SLPKDLLH 228
RESULT 11
ID NDF1_MESAU STANDARD; PRT; 355 AA.
OQ Q60430;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neurogenic differentiation factor 1 (NeuroD1) (Beta-cell E-box trans-
DE activator 2) (BETA2).
GN NEUROD1 OR NEUROD.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OC NCBI_TaxID=10036;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=95293222; PubMed=7774807;
RA Naya F.J., Stellrecht C.M.M., Tsai M.-J.;
RT "Tissue-specific regulation of the insulin gene by a novel basic
RT helix-loop-helix transcription factor.";
RL Genes Dev. 9:1009-1019(1995).
CC -1- FUNCTION: Acts as a differentiation factor during neurogenesis.
CC -1- SUBUNIT: Efficient DNA binding requires dimerization with another
CC BHLH protein. Heterodimer with E47.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- TISSUE SPECIFICITY: Most abundant in pancreatic alpha- and beta-
CC cells, less in brain and intestine.
CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U24679; AAB86518.1; ALT INIT.
CC DR InterPro; IPR001092; HLH_basic.
CC DR Pfam; PF00010; HLH; 1.
CC DR SMART; SM00353; HLH; 1.
CC DR PROSITE; PS00888; HLH; 1.
CC DR DNA-binding; Nuclear protein; Transcription regulation; Activator;
CC Neurogenesis; Developmental protein; Differentiation.
CC FT DOMAIN 58 77 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
CC FT DOMAIN 86 92 BASIC DOMAIN.
CC FT DNA BIND 101 112 BASIC DOMAIN.
CC FT DOMAIN 113 153 HELIX-LOOP-HELIX MOTIF.
CC FT DOMAIN 67 75 POLY-GLU.
CC FT DOMAIN 86 89 POLY-LYS.
CC FT SEQUENCE 355 AA; 39763 MW; F4344DFD360226B2 CRC64;
CC
CC Query Match 11.1%; Score 207.5; DB 1; Length 355;
CC Best Local Similarity 27.1%; Pred. No. 2.5e-06;
CC Matches 79; Conservative 28; Mismatches 66; Indels 119; Gaps 12;
CC
CC QY 144 RQAPSSKQVGVQ-----KQRLAANARERRRHGLNHPDOLRNVPISFNNDKLSKYE 199
CC Db : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 81 KRGRPKKKTKARLERFKLRMKANARENRHGLNADLNRKVVCPVSKTKLSKIE 140
CC
CC QY 200 TLQMAQIYINALSELQ-----QTPSG-----EQP 224
CC Db : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 141 TLRLAKNYIWAETLRSKSPDLVSFVQTLKLSQPTTNLVAGLQLNPTFLPEQNP 200
CC
CC QY 225 PPPASCKSDHHHLRTAA-----SYEGGAGNATAAGAAQAGSGSQ-----RP 276
CC Db : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 201 DMPD-----HLPTASASFPVHPYSYQSGPLSPPTGTMDSHVFQKPPPHA----- 247

```



```

QY 277 SAPASAGGYVOLDALHFTFSDSALTAMMAQKNLSPLSLQPVQ----- 324
Db 248 -----YSATLE-----PFESPUDC-----TSFSPDGLSPPLSINGNFSFKHEPS 289
QY 325 ---EEN-----SKTSPRSHRS-----DG--EFSPSHY 347
Db 290 AEFENYAFTHVPAATLAGPQSHGIFSGATAPRCEIPIDNIMSPDSSH 341

RESULT 12
NDFF6_MOUSE
ID P48986; STANDARD; PRT; 337 AA.
AC P48986;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neurogenic differentiation factor 6 (NeuroD6) (Atonal protein homolog
2) (Helix-loop-helix protein MATH-2) (MATH2) (NEX-1 protein).
GN NEUROD6 OR ATOH2 OR ATH2 OR NEX1.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/J;
RX MEDLINE=95262673; PubMed=7744035;
RA Shimizu C., Akazawa C., Nakanishi S., Kageyama R.;
RT "MATH-2, a mammalian helix-loop-helix factor structurally related to
the product of Drosophila proneural gene atonal, is specifically
expressed in the nervous system.";
RL Eur. J. Biochem. 229:239-248(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=95200803; PubMed=7545978;
RA Bartholomae A., Nave K.-A.;
RT "NEX-1: a novel brain-specific helix-loop-helix protein with
autoregulation and sustained expression in mature cortical neurons.";
RL Mech. Dev. 48:217-228(1994).
CC -!- FUNCTION: Activates E box-dependent transcription in collaboration
with E47. May be a trans-acting factor involved in the development
and maintenance of the mammalian nervous system. Transactivates
the promoter of its own gene.
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
bHLH protein.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- TISSUE SPECIFICITY: Specific to the nervous system of both embryos
and adults. Highest levels in the cortical plate of the cerebrum.
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; D44480; BAA07923.1; -.
CC EMBL; U29086; AAC14576.1; -.
CC PIR; I48682; I48682.
CC PIR; I57038; I57038.
CC MGD; MGI:106593; Neurod6.
CC InterPro; IPR001092; HLH_basic.
CC Pfam; PF00010; HLH; 1.
CC SMART; SM00353; HLH; 1.
CC PROSITE; PS50888; HLH; 1.
KW Transcription regulation; Activator; DNA-binding; Nuclear protein.
FT DOMAIN 54 63 POLY-GLU.
FT DOMAIN 80 86 BASIC DOMAIN.
FT DNA BIND 95 106 BASIC DOMAIN.
FT DOMAIN 107 147 HELIX-LOOP-HELIX MOTIF.
SQ SEQUENCE 337 AA; 38644 MW; 35C18ACD8BE1BFBA CRC64;
Query Match 11.0%; Score 205.5; DB 1; Length 337;
Best Local Similarity 29.0%; Pred. No. 3.2e-06;
Matches 71; Conservative 31; Mismatches 94; Indels 49; Gaps 7;
QY 102 RGEIVRRSSGGASSSKSPGVKVRQQLCKLKGVVVDLGCGRQRAPSSKQVNGVQ--- 157
Db 44 RKSIKRAPGEETEKEEEDDEE-----DENGLRRRLRKKTKTLRLERV 92
QY 158 KORLAANARERRRMHGLNHAFDQLNRNVPFNNDKLSKYETLQMAQIYINALSELLOT 217
Db 93 KFRQEANARERNMHLNDALDNLRLKVVPCYSKTQKLSKIETRLAKNVIWALSILRI 152
QY 218 PSGGEOP-----PPPPASCKSDHHLRTAASVEGAGNATAGAAQASGGSQ 264
Db 153 ---GKRPDLTLTFVQNLCKGLSQTFTNLVAGIQL--NARSFLMGQGGGAHHTRSPYSTFY 208
QY 265 RP-----TPPGSCRTRESAPASAGGYSVQLDALHFTFSDSALTAMMAQKNLSPLSG 317
Db 209 PPHSPELATPPGHGTLDNSKMKPNYCSAYESFYESTSPCA-----SQPFEG 258
QY 318 SILQP 322
Db 259 PLSPPP 263

RESULT 13
NGN1_RAT
ID NGN1_RAT STANDARD; PRT; 244 AA.
AC P70595;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neurogenin 1 (Neurogenic differentiation factor 3) (NeuroD3)
DE (Neurogenic basic-helix-loop-helix protein).
GN NEUROG1 OR NGN1 OR NGN OR NEUROD3.
OS Rattus norvegicus (Rat)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Embryo;
RX MEDLINE=97011137; PubMed=8858147;
RA Ma Q., Kintner C., Anderson D.J.;
RT "Identification of neurogenin, a vertebrate neuronal determination
gene.";
RL Cell 87:43-52(1996).
CC -!- FUNCTION: Appears to mediate neuronal differentiation.
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
bHLH protein.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- TISSUE SPECIFICITY: Expression restricted to the embryonic
nervous system.
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; U67777; AAC52857.1; -.
CC InterPro; IPR001092; HLH_basic.
CC Pfam; PF00010; HLH; 1.
CC SMART; SM00353; HLH; 1.
CC PROSITE; PS50888; HLH; 1.
KW DNA-binding; Nuclear protein; Transcription regulation; Activator;
KW Neurogenesis; Developmental protein; Differentiation.
FT DNA_BIND 94 105 BASIC DOMAIN.

```

FT DOMAIN 106 146 HELIX-LOOP-HELIX MOTIF.
SQ SEQUENCE 244 AA; 26182 MW; 47D0E91403143D61 CRC64;

Query Match 11.0%; Score 205; DB 1; Length 244;
Best Local Similarity 34.2%; Pred. No. 2.4e-06;
Matches 66; Conservative 24; Mismatches 69; Indels 34; Gaps 6;

QY 95 PROEVDGRLVERSSGAS-SSKSPGVKVRQELCKLKGVVVDELGCSPQAPSSQV 153
DB 38 PLASTGLSVPAARSAPTLGASNVPGQDEEQRERRRG-----RARVRSALL 87
QY 154 NGVQKQRLAANARERRRMHGLNHFADQLRNVPISFNNDKSKSYETLQMAQIYNALSE 213
DB 88 HSLRRSRVKAQDERNEMENLNAALDARSVLSPFPDDTKLTETLRFAYNYIWALAE 147
QY 214 LLQ-----TSGGSEQ-PP-----PPPASCSDHHLHRTAASYGGAGNATAAGAAQ 258
DB 148 TRLADQGLPGGARERLLPPQCVCLPGPPSPASD-----TESWGSAAASPCATV 199
QY 259 ASGGSQRTPPGS 271
DB 200 ASPLSDPSSPAS 212

RESULT 14
NDF2 RAT
ID NDF2 RAT STANDARD; PRT; 382 AA.
AC Q3689;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neurogenic differentiation factor 2 (NeuroD2) (Brain bHLH protein
DE KW8).
GN NEUROD2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Brain;
RX MEDLINE=96193685; PubMed=8605021;
RA Kume H., Maruyama K., Tomita T., Iwatsubo T., Saido T.C., Obata K.;
RT "Molecular cloning of a novel basic helix-loop-helix protein from the
rat brain."
RL Biochem. Biophys. Res. Commun. 219:526-530(1996).
RN [2]
RP CONCEPTUAL TRANSLATION.
RA Hulio C.;
RL Unpublished observations (AUG-1998).
CC -!- FUNCTION: Appears to mediate neuronal differentiation.
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- TISSUE SPECIFICITY: Detected only in neural tissue.
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC -!- CAUTION: This is a conceptual translation; two frameshifts were
CC introduced in positions 225 and 250 to extend the similarity with
CC the human and mouse orthologs.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D82868; BAAL1615.1; ALT FRAME.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00888; HLH; 1.

KW Transcription regulation; DNA-binding; Nuclear protein;
KW Differentiation.
FT DOMAIN 82 91 POLY-GLU.
FT DNAS_BIND 122 133 BASIC DOMAIN.
FT DOMAIN 134 174 HELIX-LOOP-HELIX MOTIF.
FT DOMAIN 182 285 POLY-GLY.
SQ SEQUENCE 382 AA; 41466 MW; 05214DE9BE0A8069 CRC64;

Query Match 11.0%; Score 204; DB 1; Length 382;
Best Local Similarity 23.5%; Pred. No. 4.6e-06;
Matches 93; Conservative 32; Mismatches 89; Indels 182; Gaps 16;

QY 18 DHRQPQPHLPPPPPP-----QPATLQAREHPVYPPELSILDSTDRAWLAP 67
DB 27 DEFPSDKGADPPPPPPAPGSGAPGAPATKPVSLRGEVP-----EP 68
QY 68 TLOGICTTAAAYLLHSPELGASFAAAPRDEVDGRLVRRSSGSGASSKSPGVKVRQ 127
DB 69 TLAEV-----KEEGELGGESEEEEEE-EGLE-----AEGERPKKRGPKKMT 112
QY 128 LCKLKGVVVDELGCSPQAPSSQVNGVQKQRLAANARERRRMHGLNHFADQLRNVP 187
DB 113 KARLER-----SKLRQKANARERNRMDLNALDLNRKVP 149
QY 188 SFNNDKSKSYETLQMAQIYNALSELLOT----- 217
DB 150 CYSKTQKLSKETLRLAKNYIWLSEILRSKRPDLVSYVQTLCKGLSQPTTNLVAGCLQ 209
QY 218 -----PSGGE-----QPPPPPPAS-----CKS-----DHHLRT-- 240
DB 210 LNSRNFLTEQGADGACGRFHGSGGPFAMHPYPYPCSLAGDQCOAAGGLGGGAHALRTHG 269
QY 241 -AASYE-----GGAGNATAAGAAQASGGSQRTPPGSCRTFSAPASAGYSVQLD--- 290
DB 270 YCAAYETLYAAAGGGGSPDYNSEYEG-----PLSPPLCLN-----GNFSLKQDSSP 317
QY 291 -----ALHFSITFEDSALTAMMAQKNLSPLSPGS 318
DB 318 DHEKSYHYMSYS-----AUPGS 335

RESULT 15
NGN1 MOUSE
ID NGN1 MOUSE STANDARD; PRT; 244 AA.
AC P70650;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neurogenin 1 (Neurogenic differentiation factor 3) (NeuroD3)
DE (Neurogenic basic-helix-loop-helix protein) (Helix-loop-helix protein
DE MATH-4C) (MATH4C).
GN NEUROG1 OR NGN1 OR NGN OR NEUROD3 OR ATH4C.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv;
RX MEDLINE=96413331; PubMed=8816493;
RA McCormick M.B., Tamimi R.M., Snider L., Asakura A., Bergstrom D.,
RA Tapscott S.J.;
RT "NeuroD2 and NeuroD3: distinct expression patterns and
RT transcriptional activation potentials within the neuroD gene
RT family."
RL Mol. Cell. Biol. 16:5792-5800(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv;
RX MEDLINE=97011137; PubMed=8858147;
RA Ma Q., Kintner C., Anderson D.J.;
RT "Identification of neurogenin, a vertebrate neuronal determination
RT gene.";


```

Db 52 TCAAH-AEYLLHSP--GSS-----AEGVSSASNFRKSSKS--PVKVR-LCRL 93
QY 132 KGGVVVDELGCSQRAPSSKQVNGKORLAANARERRMHGLNHAFDQLRNVIPSPFN 191
Db 94 KGAVGADE---GQRAPSSKSTNVQKORMAANARERRMHGLNHAFDQLRSVIPAFDN 150
QY 192 DKLSKYETLQMAQIYINALSELLQTPSGCEQPPPPPPASCKSDHHH-LRTAAAYEGGAGN 250
Db 151 DKLSKYETLQMAQIYINALSDLLQPGAKADPP-----NCDLLHANVLETRPRGS--- 203
QY 251 ATAAGAAQAGSGSQRPPTPGSCRTFSAPASAGYSVQLDALHFTSFEDSALTAMMAQKN 310
Db 204 -----PGVCR-----GTGVGYPYQ-----YEDGTFSFMEQDL 232
QY 311 LSP---SLGSILOPVOEENSKTSPRSHRSDGE 340
Db 233 QSPSGTSKSGS-----EASKDSPRNSRSDGE 258

RESULT 2
Q800U5 PRELIMINARY; PRT; 177 AA.
AC Q800U5;
DT 01-JUN-2003 (TEMBLrel. 24, Created)
DT 01-JUN-2003 (TEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE CATH1 (Fragment).
GN CATH1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauaria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Ebert P.J., Timmer J.R., Helms A.W., Nakada Y., Hunsaker T.L.,
RA Johnson J.E.;
RT "Control of Math1 Expression by Interactions of Zic1 and Ebf2 with the
RT Math1 Enhancer.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP467292; AAC059913.1; -.
DR InterPro; IPR001092; HLH_basic.
DR PFam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS50888; HLH 2; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 177 AA; 19170 MW; FD44269C5994F5E9 CRC64;

Query Match 28.6%; Score 532; DB 13; Length 177;
Best Local Similarity 59.6%; Pred. No. 6.1e-35;
Matches 121; Conservative 12; Mismatches 44; Indels 26; Gaps 5;

QY 152 QVNGVOKORLAANARERRMHGLNHAFDQLRNVIPSPFNNDKKLSKYETLQMAQIYINAL 211
Db 1 QVSGVOKORLAANARERRMHGLNHAFDQLRNVIPSPFNNDKKLSKYETLQMAQIYISAL 60
QY 212 SELLQTPSGCEQPPPPPPASCKSDHHHRTAAAYEGGAGNATAAGAAQAGSGSQRPPTPGS 271
Db 61 AELLHGPPA--PPEPPAKAE-----LR-GAPFE-----PPPPPPPPPPRAPPPAP 102
QY 272 CTRTFSAPASAGYSVQLDALHFTSFEDSALTAMMAQKNLSPLSGSILOPVOEENSKTS 331
Db 103 ARTFRPPAAAAGGFAAQLEPLRPPSPF-----AQKAPSPALLLGGPAPQPPERSKAS 154
QY 332 PRSHRSDGFFSPSHSYSDSDEAS 354
Db 155 PRSHRSDGFFSPSHSYSDSDEAS 177

RESULT 3
Q98939 PRELIMINARY; PRT; 161 AA.
ID Q98939
AC Q98939;

```

```

DT 01-FEB-1997 (TEMBLrel. 02, Created)
DT 01-FEB-1997 (TEMBLrel. 02, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE CATH1 (Fragment).
GN CATH1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauaria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Ben-Arie N., McCall A.E., Berkman S., Eichele G., Bellen H.J.,
RA Zoghbi H.Y.;
RT "Evolutionary conservation of sequence and expression of the bHLH
RT protein Atonal suggests a conserved role in neurogenesis.";
RL Hum. Mol. Genet. 5:1207-1216(1996).
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR EMBL; U61149; AAB41304.1; -.
DR TRANSFAC; T04545; -.
DR InterPro; IPR001092; HLH_basic.
DR PFam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS50888; HLH 2; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 161 AA; 16652 MW; BD0B73B5AA3F9F67 CRC64;

Query Match 18.1%; Score 337; DB 13; Length 161;
Best Local Similarity 61.2%; Pred. No. 2.4e-19;
Matches 79; Conservative 10; Mismatches 22; Indels 18; Gaps 5;

QY 107 RRSSGGAS---SKSPGPVKVREQLCKLKGVVVDLGCGRQRAPSSKQVNGVOKORRLA 163
Db 42 RPVAGGAARVPAGAPGE---RGAAGARGGG---GGAGPRA---QVSGVOKORRLA 89
QY 164 ANARERRRMHGLNHAFDQLRNVIPSPFNNDKKLSKYETLQMAQIYINALSELLQTPSGEQ 223
Db 90 ANARERRRMHGLNHAFDQLRNVIPSPFNNDKKLSKYETLQMAQIYISALAEHLHGPPA--- 146
QY 224 PPPPPASCK 232
Db 147 PPEPPAKAE 155

RESULT 4
Q8JHA1 PRELIMINARY; PRT; 64 AA.
AC Q8JHA1;
DT 01-OCT-2002 (TEMBLrel. 22, Created)
DT 01-OCT-2002 (TEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE BHLH transcription factor ath1 (Fragment).
OS Serinus canaria (Canary).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauaria; Aves; Neognathae; Passeriformes; Passeroidea;
OC Fringillidae; Carduelinae; Serinus.
OX NCBI_TaxID=9135;
RN [1]
RP SEQUENCE FROM N.A.
RA Huverstuhl J., Brors D., Bodmer D., Mullen L., Gleich O., Strutz J.,
RA Ryan A.F.;
RT "Expression of bHLH and class IV POU-domain transcription factors in
RT the chicken and canary inner ear.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR EMBL; AF526421; AAM89247.1; -.
DR InterPro; IPR001092; HLH_basic.
DR PFam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.

```

```

DR PROSITE; PS50888; HLH_2; 1.
FT NON TER 1
FT NON TER 64
SQ SEQUENCE 64 AA; 7439 MW; BB691E056670CA84 CRC64;

Query Match 16.6%; Score 309; DB 13; Length 64;
Best Local Similarity 95.3%; Pred. No. 1.3e-17;
Matches 61; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 153 VNGVQKORLLAANARERRRMHGLNHFADQLRNVPSPFNNDKLSKYETLQWAAQIYNALS 212
   |||||
Db 1 VSGVQKORLLAANARERRRMHGLNHFADQLRNVPSPFNNDKLSKYETLQWAAQIYISALA 60
   |||||
QY 213 ELLQ 216
   |||||
Db 61 ELLQ 64

RESULT 5
Q9NB27 PRELIMINARY; PRT; 255 AA.
AC Q9NB27;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Basic helix-loop helix transcription factor Amphineurogenin.
OS Branchiostoma floridae (Florida lancelet) (Amphioxus).
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC Branchiostoma.
OX NCBI_TaxID=77339;
RN [1]
RP SEQUENCE FROM N.A.
RA Holland L.Z., Schubert M., Holland N.D., Neuman T.;
RT "Evolutionary conservation of the presumptive neural plate markers
RT Amphioxol1/2/3 and Amphineurogenin in the invertebrate chordate
RT amphioxus.";
RL Dev. Biol. 0:0-0(2000).
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS50888; HLH_2; 1.
SQ SEQUENCE 255 AA; 28153 MW; 7EC009D66F260FB4 CRC64;

Query Match 12.5%; Score 232.5; DB 5; Length 255;
Best Local Similarity 27.2%; Pred. No. 1.1e-10;
Matches 84; Conservative 33; Mismatches 115; Indels 77; Gaps 8;

QY 38 PATLQAREHVPYPPELSLDSTDPRAWLAPTQ---GICTARAAQVLLHSPELGASEAAA 94
   |||||
Db 8 PLPRQADDMEHPGSPYSYCTDYHSLSPPTLSPSLAEVRYTDCLPSTPEY-----H 61

QY 95 PRDEVD-----GRGELV---RRSSGASSKSPGVKVRQELCKLKGVVVDLGC 143
   |||||
Db 62 PEDVMDTDSDEKAGQPRVAPTRKTENNKAETTAGTKRKSHK-----S 107

QY 144 RORAPSSKOVGVKORLLAANARERRRMHGLNHFADQLRNVPSPFNNDKLSKYETLQW 203
   |||||
Db 108 RQPKSQEAVVQKQRRKANDRENRNMHNLGALDQLREVLPDPTDKLTIKETLRF 167

QY 204 AQIYNALSLLQTPSGGPPPPPPASCKSDHHHLRTAASYEGGAGNATAAGAAQQAAGSGS 263
   |||||
Db 168 AHNYIWLSEMLKVADAGDPTVP-----MQAMPGLFGPGGLENMNIWAGGFS 217

QY 264 QRPTPGSCRTFRSAPASAGGYSVQLDALHFFSFEDSALTAMMAOKNLSPLSGSLQPV 323
   |||||
Db 218 WTP-----EQWISAPLSDPSTFPN--GSLCSEL 244

QY 324 QRENSKTS 332
   |||||
Db 245 SETDSVTYP 253

```

```

RESULT 6
Q9YH81 PRELIMINARY; PRT; 178 AA.
AC Q9YH81;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Neurogenin 1.
GN NGN1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21064448; PubMed=11124117;
RA Matter-Sadzinski L., Matter J.M., Ong M.T., Hernandez J., Ballivet M.;
RT "Specification of neurotransmitter receptor identity in developing
RT retina: the chick ATH5 promoter integrates the positive and negative
RT effects of several BHLH proteins.";
RL Development 128:217-231(2001).
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR EMBL; AJ012660; CAA10106.1; -.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS50888; HLH_2; 1.
SQ SEQUENCE 178 AA; 19058 MW; 7EBCC1CD1ADA5633 CRC64;

Query Match 12.3%; Score 228.5; DB 13; Length 178;
Best Local Similarity 34.7%; Pred. No. 1.4e-10;
Matches 77; Conservative 26; Mismatches 54; Indels 65; Gaps 11;

QY 109 SSGGASSKSPGVKVRQELCKLKGVVVDLGCGRAPSKQVGVKORLLAANARE 168
   |||||
Db 7 SSGGV--SEPGAPERR---RRRG-----RARARTALHLLTKRSRRVKANDRE 51

QY 169 RRRMHGLNHFADQLRNVPSPFNNDKLSKYETLQWAAQIYNALSLLQTPSGEQ--PPP 226
   |||||
Db 52 RNRMHNLNALDELRSVLPTPDPTDKLTIKETLRFAYNYIWLSETLRL--AEQCLPPP 108

QY 227 P-----PASCSDHHHLRTAASYEGGAGNATAAGAAQQAAGSGSORPTPGSCRTFRSAP 279
   |||||
Db 109 PAFRGPPAPPSPGSD-----AGSWLSGSGS---PAAPSLC----- 139

QY 280 ASAGYSVQLDALHFFSFEDSALT---AMMAOKNLSPLSGS 318
   |||||
Db 140 ASAGSPSP-----ATSEDCGVVPSDALRPRGLPPAPGA 175

RESULT 7
Q9PWP1 PRELIMINARY; PRT; 179 AA.
AC Q9PWP1;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Neurogenin 1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99180576; PubMed=10079233;
RA Perez S.E., Rebelo S., Anderson D.J.;
RT "Early specification of sensory neuron fate revealed by expression and
RT function of neurogenins in the chick embryo.";

```

Qy	92	AAAPRDEVDRGELVRSSSGASSKSGPGVKVREQLCKLKGVVVDBLQSGSRQAPSSK	151
Db	35	GEDEDEBDRPR--RLQEGARRAGRQRPRA-----ARTAE	70
Qy	152	QVNGVQQRLLAANARERRRMEGINHAFDQLRNVIPSENNDKLSYETLQWAIYIWA	211
Db	71	TAQIKKSRRLKANRRERNMHNLNAALDALRDVLPTPEDAKUTKIEITLFAHNYIWA	130
Qy	212	SELLQTPS-----GGEQPPPPSPASCKSDHHHLRTAASVEGGAGNATAAG-----A	256
Db	131	TETLRLAGARLGAADAAPCAAAAGSP---SPASSWGGSPAPSPASPYACTLSFGSPA	187
Qy	257	QQASGGGQRPTPPGSCSTRFSAP	279
Db	188	GSASDAEHWPMPPRG-----RFAPP	206

```

RESULT 9
Q9PWP2
ID Q9PWP2 PRELIMINARY; PRT; 211 AA.
AC Q9PWP2;
DT 01-MAY-2000 (TRMBLrel. 13, Created)
DT 01-MAY-2000 (TRMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TRMBLrel. 23, Last annotation update)
DE Neurogenin 2 (Fragment).
OS Gallus Gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OC NCBI TaxID=9031;

```

```

RN      SEQUENCE FROM N. A.
RP      MEDLINE=99180576; PubMed=10079233;
RX      Perez S.E., Rebelo S., Anderson D.J.;
RT      "Early specification of sensory neuron fate revealed by expression and
RT      function of neurogenins in the chick embryo.";
RL      Development 126:1715-1728(1999).
CC      C-1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC      TRANSCRIPTION FACTORS.
DR      EMBL: AF123884; AAD22060.1; -.
DR      InterPro: IPR001092; HLH_basic.
DR      Pfam: PF00010; HLH; 1.
DR      SMART: SM00353; HLH; 1.
DR      PROSITE: PS50888; HLH 2; 1.
FT      NON_TER
SQ      SEQUENCE      211 AA; 22278 MW; 30ACDB332CA6356D CRC64;

```

```

144 ADAAPCAAGGSP---SPASSWGGGASPAFSAFYACTLSPGSPAGSADAEHWPPRG- 199
Db

272 CRTRESAP 279
Qy

200 ---RFAPP 204
Db

```

RESULT 10


```

O9W6C7
ID Q9W6C7 PRELIMINARY; PRT; 316 AA.
AC Q9W6C7;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Basic helix-loop-helix transcription factor Ndr1b.
GN ATOH2B OR NDR1B.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OC NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9249195; PubMed=10235116;
RA Liao J., He J., Yan T., Korzh V., Gong Z.;
RT "A class of neuroD-related basic helix-loop-helix transcription
RT factors expressed in developing central nervous system in zebrafish.";
RL DNA Cell Biol. 18:333-344(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Wang X., Korzh V., Gong Z.;
RT "The functional specificity of NeuroD protein is defined by a single
RT amino acid residue (N11) in the basic domain.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR EMBL; AF115773; AAD23442.2; -.
DR ZFIN; ZDB-GENE-010608-2; atch2b.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00038; HLH 1; 1.
DR PROSITE; PS50888; HLH 2; 1.
DR NCBI_TaxID=7955;
SQ SEQUENCE 316 AA; 36444 MW; 228AEDC5369AE60 CRC64;

Query Match 11.9%; Score 222.5; DB 13; Length 316;
Best Local Similarity 27.2%; Pred. No. 9e-10;
Matches 81; Conservative 40; Mismatches 96; Indels 81; Gaps 12;

OY 90 SEAAAPRDEVDGRLVRRSGGASSKSPGVKVRQCLKLGKGVVVDLGL 141
DB 15 SQFGPRSR-----VRKTSGHSAAPSRQRTTTRTCRRRREE-----DENG 57
OY 142 CSRQAPSSQVNG---VQQRRLAANARERRRMHGLNHPADQLRNVIP 198
DB 58 LPKXGPRKKSEGRGDRVQMRQANARERRRMHGLNDALESRLKVPYCKTKLSKI 117
OY 199 ETLQMAQIYINALSELLQTPSGGQPPP---PPASCKSDHHLRTAASVEGGAGNAT--- 252
DB 118 ETLRLAKNYIWALET---SAGRPDLIAFVQTLCK-----GLSQPTTWL 160
OY 253 AAGAQAQSGSORPTPGSCRTFRFASAGYSVOLDALH----- 293
DB 161 VAGCLQLNARNFLTDHNGTCRS- LAGPRITPCTHTRPKCHAHRPQLWDARKQTPRYN 219
OY 294 -----FSTFEDSALTAMMAQKNLSPLSPGSILOPQVQENSKTSPRSRSGHSPSHY 347
DB 220 YVASYESYDASA-----SPSSSPFDGQMSNPINYNGLSLKHKHDEQVEYSKNCHY 271

RESULT 11
Q8N416
ID Q8N416 PRELIMINARY; PRT; 272 AA.
AC Q8N416;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

O9W6C7
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR EMBL; BC036847; AAH36847.1; -.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS50888; HLH 2; 1.
DR Hypothetical protein.
KW NCBI_TaxID=9606;
SQ SEQUENCE 272 AA; 28621 MW; 57AB37CC4A9E15FC CRC64;

Query Match 11.9%; Score 222; DB 4; Length 272;
Best Local Similarity 30.4%; Pred. No. 8.1e-10;
Matches 90; Conservative 27; Mismatches 89; Indels 90; Gaps 14;

OY 84 SPDLGA-----SEAAAPRDEVDGRLVRRSGGASSKSPGVKVRQCLKLGK--GVV 136
DB 24 SPALALTLPLSSADEEEEPFGAGGARRORGAAGQAGGVAAGAGCEGPALLGLV 83
OY 137 VDELGC-----SRQAPS-----SKQVGVQKORRLAANARERRRMHGLNHPADQLRNVIP 187
DB 84 HD---CKRRPSEARAVSRGAKTAETVQRIKTRRLKANNRRNRHNLNALDALREVLP 140
OY 188 SPNDKLSKYETLQMAQIYINALSELLQTPSGGQPPPAPASCKSDHHLRTAASVEGG 247
DB 141 TFPEDAKLTKETLRFHNYIWALETTLR-----LADHCGG 176
OY 248 AGNAT-----AAGAQAQSGSORPTPGSCRTFRFASAGYSVOLDAL 292
DB 177 GGGGLPGALFSEAVLLSPGASALSSGDSFSPASTWSC-TNSAPSS----- 225
OY 293 HFSTFEDSALTAMMAQKNLSPLSP- GSIL---QPVOENSKTSPRSRSGHSPSH 344
DB 226 -VSSNSTSPYSC-----TILSPASPGSDMDYWPQ-----PPDKHR-----YAPH 264

RESULT 12
O9PTR9
ID O9PTR9 PRELIMINARY; PRT; 213 AA.
AC O9PTR9;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Neurogenin 2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White Leghorn; TISSUE=Brain;
RA Van R.-T., Li C.-M., Wang S.-Z.;
RT "Neurogenin 2 acts upstream of NEUROD during rpe transdifferentiation
RT into a photoreceptor cell.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR EMBL; AF109014; AAF21904.1; -.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS50888; HLH 2; 1.
DR NCBI_TaxID=9031;
SQ SEQUENCE 213 AA; 22497 MW; BF6EC92A4DA5PB26 CRC64;

Query Match 11.8%; Score 220.5; DB 13; Length 213;
Best Local Similarity 29.3%; Pred. No. 7.8e-10;
Matches 77; Conservative 24; Mismatches 89; Indels 73; Gaps 9;

```

QY	279	PASAGGYSVQLDALHSTFSDSALTAMMAQK-----NLSPS-----LPQSLQPVOE	325
Db	222	-----YGMESALLHLKGYGPPYENSPPNRCSSGTTPYDGLTFPLSI	266
QY	326	ENS---KTSPPSRHSRDGEFSPH-SHYSDS	350
Db	267	SGNFALKQEPSPREARNFTPHPTHYISS	295

	PRELIMINARY;	PRT; 347 AA.
RESULT 14		
Q8JH35		
ID Q8JH35		
AC Q8JH35		
DT 01-OCT-2002	(TRMBLrel. 22, Created)	
DI 01-OCT-2002	(TRMBLrel. 22, Last sequence update)	
DE 01-MAR-2003	(TRMBLrel. 23, Last annotation update)	
DE Atonal-like protein_3		

CC Cyprinidae; Danio.
 CC NCBI_TaxID=7955;
 CC [1]
 RN SEQUENCE FROM N.A.
 RP Wang X., Korzh V., Gong Z.;
 RA "zath3 plays roles in the development of zebrafish neurons and glia
 CC cells.";
 RT Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 RL
 CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
 CC TRANSCRIPTION FACTORS.

SQ SEQUENCE 347 AA; 38317 MW; 0968893797AEE4B2 CRC64;

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 20:34:21 ; Search time 167.088 Seconds
(without alignments)
598.618 Million cell updates/sec

Title: US-09-980-381A-58
 Perfect score: 1862
 Sequence: 1 MSRLHAEWEAEVKELGDHH.....HRSDGFSPHSHYSDSDEAS 354

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database : A_Geneseq_29Jan04.*
1: Genesexp1980s.*
2: Genesexp1990s.*
3: Genesexp2000s.*
4: Genesexp2001s.*
5: Genesexp2002s.*
6: Genesexp2003as.*
7: Genesexp2003bs.*
8: Genesexp2004s.*
```

SUMMARIES

Result No.	Score	Query %		DB	ID	Description
		Match	Length			
1	1862	100.0	354	4	AAb603375	Aab603375 Human ato
2	1862	100.0	354	4	AAb60349	Aab60349 Human ato
3	1862	100.0	354	6	AAb603945	Aao303945 Human ato
4	1862	100.0	354	6	AAb6030920	Aao3030920 Human ato
5	1627.5	87.4	351	4	AAb603353	Aab603353 Mouse ato
6	1627.5	87.4	351	4	AAb603369	Aab603369 Mouse ato
7	1627.5	87.4	351	6	AAb603939	Aao303939 Mouse ato
8	1627.5	87.4	351	6	AAb603924	Aao303924 Mouse ato
9	582	31.3	266	4	AAb603366	Aab603366 zebrafish
10	582	31.3	266	6	AAb603936	Aao303936 zebra fis
11	537	28.8	259	6	AAb603949	Aao303949 Frog aton
12	337	18.1	161	4	AAb603376	Aab603376 Chicken a
13	337	18.1	161	6	AAb603946	Aao303946 Chicken a
14	228.5	12.3	178	4	AAb603356	Aab603356 Chicken a
15	228.5	12.3	178	6	AAb6030926	Aao30926 Chicken a
16	228.5	12.3	198	4	ABb71919	Abb71919 Drosophil
17	224.5	12.1	213	4	AAb603357	Aab603357 Chicken a
18	224.5	12.1	213	6	AAb603927	Aao303927 Chicken a
19	222	11.9	250	5	AAe29281	AAe29281 Human neu
20	222	11.9	272	7	ADc31349	Adc31349 Human nov
21	218	11.7	263	4	AAb603351	Aab603351 Mouse ato
22	218	11.7	263	4	AAb603365	Aab603365 Mouse ato
23	218	11.7	263	6	AAb603935	Aao303935 Mouse ato
24	218	11.7	263	6	AAb603922	Aao303922 Mouse ato
25	217.5	11.7	195	4	ABb62789	Abb62789 Drosophil

ALIGNMENTS

RESULT 1	
AAB60375	
ID	AAB60375 standard; protein; 354 AA.
XX	
XX	AAB60375;
AC	
XX	
DT	24-APR-2001 (first entry)
XX	
DE	Human atonal homologue 1 (Hath1) protein. SEQ ID NO:58.

26	217	11.7	151	4	AAB60364	Chicken a
27	217	11.7	151	6	AAB030934	Chicken a
28	217	11.7	263	2	AAW54946	Mouse neu
29	217	11.7	263	3	AAY70569	Murine neu
30	217	11.7	263	4	AAB60362	Mouse neu
31	217	11.7	263	4	AAC030932	Mouse ato
32	217	11.7	352	2	AAR77505	Frog neur
33	217	11.5	352	2	AAR77505	Frog neur
34	214	11.5	352	2	AAR22437	Frog neur
35	214	11.5	352	2	AAG71010	Xenopus nr
36	213	11.5	352	6	ABG71796	Frog bHLH
37	213	11.4	312	4	ABB63330	Drosophill
38	213	11.4	312	4	AAB60378	Drosophill
39	213	11.4	312	6	AAC030948	Drosophill
40	212.5	11.4	189	6	AAB60355	Drosophill
41	212.5	11.4	189	6	AAC030925	Drosophill
42	210.5	11.3	381	2	AAR22439	Human neu
43	210.5	11.3	381	2	AAG71015	Human neu
44	210.5	11.3	381	6	ABG72001	Human bHL
45	210	11.3	382	3	AAB14347	Human neu

CC gene is replaced by a heterologous nucleic acid sequence, thus
 CC inactivating the atonal-associated allele. The nucleic acids or proteins
 CC may be used in a method of treating an animal for hearing impairment,
 CC joint disease, balance disorders, abnormal cell proliferation, or other
 CC disease related to loss of a functional atonal-associated nucleic acid or
 CC protein. They may particularly be used to treat an animal with a
 CC deficiency in cerebellar granule neurons or their precursors, and may
 CC also be used in promoting mechanoreceptive cell growth and generating
 CC hair cells. The present sequence represents an atonal-associated amino
 CC acid sequence referred to in the invention. Note: The present sequence is
 CC not shown in the specification, but was obtained from GenBank
 XX
 SQ Sequence 354 AA;

Query Match 100.0%; Score 1862; DB 4; Length 354;
 Best Local Similarity 100.0%; Pred. No. 4.2e-151;
 Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSRLHAEWAELGDKHHRQPHLPPPPPPPPATLQAREHVPVPELSDSTD 60
 Db 1 MSRLHAEWAELGDKHHRQPHLPPPPPPPPATLQAREHVPVPELSDSTD 60
 QY 61 PRAWLPTLQIGICTARAAQYLLHSPGLGASEAAAPRDEVDGRLVRRSGGASSKSPG 120
 Db 61 PRAWLPTLQIGICTARAAQYLLHSPGLGASEAAAPRDEVDGRLVRRSGGASSKSPG 120
 QY 121 PVKVRQOLCKLGGVVVDLGCSPORAPSSKQVNGVKORRLAANARERRMGLNHAFD 180
 Db 121 PVKVRQOLCKLGGVVVDLGCSPORAPSSKQVNGVKORRLAANARERRMGLNHAFD 180
 QY 181 QLRNVIPSFNNDKLSKYETLQMAQIYINALSELLQTPSGGEQPPPPASCKSDHHHLRT 240
 Db 181 QLRNVIPSFNNDKLSKYETLQMAQIYINALSELLQTPSGGEQPPPPASCKSDHHHLRT 240
 QY 241 AASYEGAGNATAAGAAQAGSGGSGRPTPGSCRTFRSAPASAGYSVQLDALHFSTPDS 300
 Db 241 AASYEGAGNATAAGAAQAGSGGSGRPTPGSCRTFRSAPASAGYSVQLDALHFSTPDS 300
 QY 301 ALTAMMAQKNLSPLSGSILQPVOEENSKTSPRSHRSDGFEFSPHSYSDSDEAS 354
 Db 301 ALTAMMAQKNLSPLSGSILQPVOEENSKTSPRSHRSDGFEFSPHSYSDSDEAS 354

RESULT 2
 AAB60349
 ID AAB60349 standard; protein; 354 AA.
 XX AAB60349;
 AC AAB60349;
 XX
 DT 24-APR-2001 (first entry)
 XX
 DE Human atonal homologue 1 (ATOH1, Hath1) protein, SEQ ID NO:2.
 XX
 KW Atonal; homologue; orthologue; atonal-associated protein; deafness;
 KW hearing impairment; vestibular effect; balance disorder; osteoarthritis;
 KW cellular proliferation; cerebellar granule neuron; gene therapy;
 KW mechanoreceptive cell growth; auditory; osteopathic; cytostatic;
 KW transgenic animal.
 XX
 OS Homo sapiens.
 XX
 FN WO200073764-A2.
 XX
 PD 07-DEC-2000.
 XX
 PF 01-JUN-2000; 2000WO-US015410.
 XX
 PR 01-JUN-1999; 99US-0137060P.
 PR 19-JAN-2000; 2000US-0176993P.
 XX
 XX (BAYU) BAYLOR COLLEGE MEDICINE.
 PA
 XX Zoghbi HY, Bellen H, Birmingham N, Hassan B, Ben-Arie N;
 FI

XX WPI: 2001-032190/04.
 DR N-PSDB; AAF27252.
 XX
 PT Therapeutic use of atonal-associated nucleic acids or amino acids, or any
 PT of its homologs or orthologs, for the treatment of e.g. deafness,
 PT osteoarthritis and abnormal cell proliferation.
 XX
 PS Disclosure; Page; 142pp; English.
 XX
 CC The invention relates to the use of atonal-associated nucleic acid or
 CC amino acid sequence, or any of its homologs or orthologues as
 CC therapeutic agents for the treatment of deafness, partial hearing loss,
 CC vestibular effects due to damage or loss of inner hair cells,
 CC osteoarthritis and abnormal cell proliferation. The invention also
 CC encompasses methods of screening for compounds which affect the
 CC expression of an atonal-associated nucleic acid sequence in an animal,
 CC and a transgenic animal in which an allele of a native atonal-associated
 CC gene is replaced by a heterologous nucleic acid sequence, thus
 CC inactivating the atonal-associated allele. The nucleic acids or proteins
 CC may be used in a method of treating an animal for hearing impairment,
 CC joint disease, balance disorders, abnormal cell proliferation, or other
 CC disease related to loss of a functional atonal-associated nucleic acid or
 CC protein. They may particularly be used to treat an animal with a
 CC deficiency in cerebellar granule neurons or their precursors, and may
 CC also be used in promoting mechanoreceptive cell growth and generating
 CC hair cells. The present sequence represents an atonal-associated amino
 CC acid sequence referred to in the invention. Note: The present sequence is
 CC not shown in the specification, but was obtained from GenBank
 XX
 SQ Sequence 354 AA;

Query Match 100.0%; Score 1862; DB 4; Length 354;
 Best Local Similarity 100.0%; Pred. No. 4.2e-151;
 Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSRLHAEWAELGDKHHRQPHLPPPPPPPPATLQAREHVPVPELSDSTD 60
 Db 1 MSRLHAEWAELGDKHHRQPHLPPPPPPPPATLQAREHVPVPELSDSTD 60
 QY 61 PRAWLPTLQIGICTARAAQYLLHSPGLGASEAAAPRDEVDGRLVRRSGGASSKSPG 120
 Db 61 PRAWLPTLQIGICTARAAQYLLHSPGLGASEAAAPRDEVDGRLVRRSGGASSKSPG 120
 QY 121 PVKVRQOLCKLGGVVVDLGCSPORAPSSKQVNGVKORRLAANARERRMGLNHAFD 180
 Db 121 PVKVRQOLCKLGGVVVDLGCSPORAPSSKQVNGVKORRLAANARERRMGLNHAFD 180
 QY 181 QLRNVIPSFNNDKLSKYETLQMAQIYINALSELLQTPSGGEQPPPPASCKSDHHHLRT 240
 Db 181 QLRNVIPSFNNDKLSKYETLQMAQIYINALSELLQTPSGGEQPPPPASCKSDHHHLRT 240
 QY 241 AASYEGAGNATAAGAAQAGSGGSGRPTPGSCRTFRSAPASAGYSVQLDALHFSTPDS 300
 Db 241 AASYEGAGNATAAGAAQAGSGGSGRPTPGSCRTFRSAPASAGYSVQLDALHFSTPDS 300
 QY 301 ALTAMMAQKNLSPLSGSILQPVOEENSKTSPRSHRSDGFEFSPHSYSDSDEAS 354
 Db 301 ALTAMMAQKNLSPLSGSILQPVOEENSKTSPRSHRSDGFEFSPHSYSDSDEAS 354

RESULT 3
 AAO30945
 ID AAO30945 standard; protein; 354 AA.
 XX AAO30945;
 AC AAO30945;
 XX
 DT 22-SEP-2003 (first entry)
 XX
 DE Human atonal homologue 1 (Hath1) protein #2.
 XX
 KW Atonal-associated protein; gastrointestinal; damaged intestinal tissue;
 KW cancer; inflammatory bowel disease; irritable bowel syndrome; infection;
 KW

KW necrotising enterocolitis; hearing impairment; abnormal proliferation;
 KW joint disease; gene therapy; imbalance disorder; atonal homologue 1;
 KW Hathi; human.
 XX
 OS Homo sapiens.
 XX
 PN WO2003047532-A2.
 XX
 XX 12-JUN-2003.
 XX
 PD 03-DEC-2002; 2002WO-US041458.
 XX
 PF 05-DEC-2001; 2001US-00004717.
 XX
 PR (BAYU) BAYLOR COLLEGE MEDICINE.
 XX
 PA
 XX
 PI Zoghbi HY, Qi Y;
 XX
 XX WPI; 2003-505253/47.
 DR N-PSDB; AAL62230.
 XX
 XX New compositions comprising at least one stem cell which is up- or down-
 PT regulated for expression of an atonal-associated sequence, and at least
 PT one regulatory factor, useful for treating a gastrointestinal condition,
 PT e.g. cancer.
 XX
 XX Claim 33; Page 151-152; 157pp; English.
 PS
 XX The invention relates to pharmaceutical compositions and methods for the
 CC therapeutic use of an atonal-associated sequence. The composition
 CC comprises of a stem cell which is up- or down-regulated for expression of
 CC an atonal-associated sequence and a regulatory factor. Atonal-associated
 CC amino acid or nucleic acid sequence is used to treat a gastrointestinal
 CC condition including cancer, damaged intestinal tissue, inflammatory bowel
 CC disease, irritable bowel syndrome, infection or necrotising enterocolitis.
 CC They are also useful for promoting mechanoreceptive cell growth, for
 CC generating hair cells, for treating hearing impairment or imbalance
 CC disorder, joint disease, abnormal proliferation (e.g. cancer), or a
 CC disease that is a result or loss of functional atonal-associated
 CC sequences. Atonal-associated sequences are also used in gene therapy. The
 CC present sequence is human atonal homologue 1 (Hathi) protein. This
 CC sequence is used to illustrate the method of the invention
 XX
 XX Sequence 354 AA;
 PS
 XX Query Match 100.0%; Score 1862; DB 6; Length 354;
 Best Local Similarity 100.0%; Pred. No. 4.2e-151;
 Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSRLHAEWAENVKELGDHHRQPQPHHLPQPPPPQPPATLQAREHPVYPPELSDSTD 60
 DB 1 MSRLHAEWAENVKELGDHHRQPQPHHLPQPPPPQPPATLQAREHPVYPPELSDSTD 60
 QY 61 PRAWLAPTQIGICTARAAQYLLHSPELGASAAAAPRDEVDGRGELVRRSSGASSKSPG 120
 DB 61 PRAWLAPTQIGICTARAAQYLLHSPELGASAAAAPRDEVDGRGELVRRSSGASSKSPG 120
 QY 121 PVKVRQLCKLKGVVVDDELGCSRQAPSSKQVNGVKORRLAANARERRMHLNHAFD 180
 DB 121 PVKVRQLCKLKGVVVDDELGCSRQAPSSKQVNGVKORRLAANARERRMHLNHAFD 180
 QY 181 QLRNVTPSFNNDDKLSKYETLQMAQIYINALSELLQTPSGGEGOPPPPPASCKSDHHHLRT 240
 DB 181 QLRNVTPSFNNDDKLSKYETLQMAQIYINALSELLQTPSGGEGOPPPPPASCKSDHHHLRT 240
 QY 241 AASYEGAGNATAAGQAQSGGSGRTPPGSCRTFSAPASAGYSVQDLALHFSFEDS 300
 DB 241 AASYEGAGNATAAGQAQSGGSGRTPPGSCRTFSAPASAGYSVQDLALHFSFEDS 300
 QY 301 ALTAMMAQKNLSPSLPGSILQPVQENSKTSPRSHRSDGFEFSPHSHYSDEAS 354
 DB 301 ALTAMMAQKNLSPSLPGSILQPVQENSKTSPRSHRSDGFEFSPHSHYSDEAS 354

RESULT 4
 AAO30920
 ID AAO30920 standard; protein; 354 AA.
 XX
 AC AAO30920;
 XX
 DT 22-SEP-2003 (first entry)
 XX
 DE Human atonal homologue 1 (Hathi) protein #1.
 XX
 XX Atonal-associated protein; gastrointestinal; damaged intestinal tissue;
 KW cancer; inflammatory bowel disease; irritable bowel syndrome; infection;
 KW necrotising enterocolitis; hearing impairment; abnormal proliferation;
 KW joint disease; gene therapy; imbalance disorder; atonal homologue 1;
 KW Hathi; human.
 XX
 OS Homo sapiens.
 XX
 PN WO2003047532-A2.
 XX
 XX 12-JUN-2003.
 PD 03-DEC-2002; 2002WO-US041458.
 XX
 PF 05-DEC-2001; 2001US-00004717.
 XX
 PR (BAYU) BAYLOR COLLEGE MEDICINE.
 XX
 PA
 XX
 PI Zoghbi HY, Qi Y;
 XX
 XX WPI; 2003-505253/47.
 DR N-PSDB; AAL62199.
 XX
 XX New compositions comprising at least one stem cell which is up- or down-
 PT regulated for expression of an atonal-associated sequence, and at least
 PT one regulatory factor, useful for treating a gastrointestinal condition,
 PT e.g. cancer.
 XX
 XX Disclosure; Page 115-117; 157pp; English.
 PS
 XX The invention relates to pharmaceutical compositions and methods for the
 CC therapeutic use of an atonal-associated sequence. The composition
 CC comprises of a stem cell which is up- or down-regulated for expression of
 CC an atonal-associated sequence and a regulatory factor. Atonal-associated
 CC amino acid or nucleic acid sequence is used to treat a gastrointestinal
 CC condition including cancer, damaged intestinal tissue, inflammatory bowel
 CC disease, irritable bowel syndrome, infection or necrotising enterocolitis.
 CC They are also useful for promoting mechanoreceptive cell growth, for
 CC generating hair cells, for treating hearing impairment or imbalance
 CC disorder, joint disease, abnormal proliferation (e.g. cancer), or a
 CC disease that is a result or loss of functional atonal-associated
 CC sequences. Atonal-associated sequences are also used in gene therapy. The
 CC present sequence is human atonal homologue 1 (Hathi) protein. This
 CC sequence is used to illustrate the method of the invention
 XX
 XX Sequence 354 AA;
 PS
 XX Query Match 100.0%; Score 1862; DB 6; Length 354;
 Best Local Similarity 100.0%; Pred. No. 4.2e-151;
 Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSRLHAEWAENVKELGDHHRQPQPHHLPQPPPPQPPATLQAREHPVYPPELSDSTD 60
 DB 1 MSRLHAEWAENVKELGDHHRQPQPHHLPQPPPPQPPATLQAREHPVYPPELSDSTD 60
 QY 61 PRAWLAPTQIGICTARAAQYLLHSPELGASAAAAPRDEVDGRGELVRRSSGASSKSPG 120
 DB 61 PRAWLAPTQIGICTARAAQYLLHSPELGASAAAAPRDEVDGRGELVRRSSGASSKSPG 120
 QY 121 PVKVRQLCKLKGVVVDDELGCSRQAPSSKQVNGVKORRLAANARERRMHLNHAFD 180
 DB 121 PVKVRQLCKLKGVVVDDELGCSRQAPSSKQVNGVKORRLAANARERRMHLNHAFD 180

QY 181 QLRNVIPSFNNDDKLSKYETLQMAQIYINALSELLQTPSGGQPPPPPPASCKSDHHHLRT 240
 Db 181 QLRNVIPSFNNDDKLSKYETLQMAQIYINALSELLQTPSGGQPPPPPPASCKSDHHHLRT 240
 QY 241 AASVEGGAGNATAAGAAQAGSQRPTPGSCRTFRFSAPASAGYSVQLDALHFSTFEDS 300
 Db 241 AASVEGGAGNATAAGAAQAGSQRPTPGSCRTFRFSAPASAGYSVQLDALHFSTFEDS 300
 QY 301 ALTAMMAQKNLSPSLPGSILOPVOEENSKTSPRSHRSDGFSPSHYSDSDEAS 354
 Db 301 ALTAMMAQKNLSPSLPGSILOPVOEENSKTSPRSHRSDGFSPSHYSDSDEAS 354

RESULT 5
 AAB60353
 ID AAB60353 standard; protein; 351 AA.
 XX
 AC AAB60353;
 DT 24-APR-2001 (first entry)
 XX
 DE Mouse atonal homologue 1 (ATOH1, Math1) protein, SEQ ID NO:11.
 XX
 KW Atonal; homologue; orthologue; atonal-associated protein; deafness;
 KW hearing impairment; vestibular effect; balance disorder; osteoarthritis;
 KW cellular proliferation; cerebellar granule neuron; gene therapy;
 KW mechanoreceptive cell growth; auditory; osteopathic; cyostatic;
 KW transgenic animal.
 XX
 OS Mus musculus.
 XX
 PN WO200073764-A2.
 XX
 PD 07-DEC-2000.
 XX
 PF 01-JUN-2000; 2000WO-US015410.
 XX
 PR 01-JUN-1999; 99US-0137060P.
 PR 19-JAN-2000; 2000US-0176993P.
 XX
 PA (BAYU) BAYLOR COLLEGE MEDICINE.
 XX
 PI Zoghbi HY, Bellen H, Birmingham N, Hassan B, Ben-Arie N;
 XX
 DR WPI; 2001-032190/04.
 DR N-PSDB; AAF27257.
 XX
 PT Therapeutic use of atonal-associated nucleic acids or amino acids, or any
 PT of its homologs or orthologs, for the treatment of e.g. deafness,
 PT osteoarthritis and abnormal cell proliferation.
 XX
 PS Disclosure; Page; 142pp; English.
 XX
 CC The invention relates to the use of atonal-associated nucleic acid or amino acid sequence, or any of its homologs or orthologues as therapeutic agents for the treatment of deafness, partial hearing loss, vestibular effects due to damage or loss of inner hair cells, osteoarthritis and abnormal cell proliferation. The invention also encompasses methods of screening for compounds which affect the expression of an atonal-associated nucleic acid sequence in an animal, and a transgenic animal in which an allele of a native atonal-associated gene is replaced by a heterologous nucleic acid sequence, thus inactivating the atonal-associated allele. The nucleic acids or proteins may be used in a method of treating an animal for hearing impairment. CC joint disease, balance disorders, abnormal cell proliferation, or other disease related to loss of a functional atonal-associated nucleic acid or protein. They may particularly be used to treat an animal with a deficiency in cerebellar granule neurons or their precursors, and may also be used in promoting mechanoreceptive cell growth and generating hair cells. The present sequence represents an atonal-associated amino acid sequence referred to in the invention. Note: The present sequence is not shown in the specification, but was obtained from GenBank

XX
 SQ Sequence 351 AA;
 Query Match 87.4%; Score 1627.5; DB 4; Length 351;
 Best Local Similarity 88.4%; Pred. No. 5.4e-131;
 Matches 313; Conservative 11; Mismatches 27; Indels 3; Gaps 2;
 QY 1 MSRLHAEAEVAEVELGDDHRRQPOPHLPQPPPPPPATLQAREHPVYPELSILDSTD 60
 Db 1 MSRLHAEAEVAEVELGDDHRRQPOPHV--PPLTPQPPATLQARDLPVYPAELSLDSTD 58
 QY 61 PRAWLPTLOQICTARAAQYLLHSPELGASEAAAPRDEVQGRGELVRRSSGAGSSXSPG 120
 Db 59 PRAWLTPTLOGLCTARAAQYLLHSPELGASEAAAPRDEADSQGLVLR--SCGGLSKSPG 117
 QY 121 PVKVRQQLCKLGGVVVDELGCSRQAPSSKQVNGVQKRRRLAANARERRMHGLNHAFD 180
 Db 118 PVKVRQQLCKLGGVVVDELGCSRQAPSSKQVNGVQKRRRLAANARERRMHGLNHAFD 177
 QY 181 QLRNVIPSFNNDDKLSKYETLQMAQIYINALSELLQTPSGGQPPPPPPASCKSDHHHLRT 240
 Db 178 QLRNVIPSFNNDDKLSKYETLQMAQIYINALSELLQTPNVGGEQPPPTASCKNDHHHLRT 237
 QY 241 AASVEGGAGNATAAGAAQAGSQRPTPGSCRTFRFSAPASAGYSVQLDALHFSTFEDS 300
 Db 238 ASSVEGGAGASAVAGAQPAPGGGPRPTPGPCRTFRFSGPASSGGYSVQLDALHFPFEDR 297
 QY 301 ALTAMMAQKNLSPSLPGSILOPVOEENSKTSPRSHRSDGFSPSHYSDSDEAS 354
 Db 298 ALTAMMAQKDLSPSLPGGILQPVQEDNSKTSPRSHRSDGFSPSHYSDSDEAS 351

RESULT 6
 AAB60369
 ID AAB60369 standard; protein; 351 AA.
 XX
 AC AAB60369;
 DT 24-APR-2001 (first entry)
 XX
 DE Mouse atonal homologue 1 (Math-1) protein, SEQ ID NO:46.
 XX
 KW Atonal; homologue; orthologue; atonal-associated protein; deafness;
 KW hearing impairment; vestibular effect; balance disorder; osteoarthritis;
 KW cellular proliferation; cerebellar granule neuron; gene therapy;
 KW mechanoreceptive cell growth; auditory; osteopathic; cyostatic;
 KW transgenic animal.
 XX
 OS Mus musculus.
 XX
 PN WO200073764-A2.
 XX
 PD 07-DEC-2000.
 XX
 PF 01-JUN-2000; 2000WO-US015410.
 XX
 PR 01-JUN-1999; 99US-0137060P.
 PR 19-JAN-2000; 2000US-0176993P.
 XX
 PA (BAYU) BAYLOR COLLEGE MEDICINE.
 XX
 PI Zoghbi HY, Bellen H, Birmingham N, Hassan B, Ben-Arie N;
 XX
 DR WPI; 2001-032190/04.
 DR N-PSDB; AAF27277.
 XX
 PT Therapeutic use of atonal-associated nucleic acids or amino acids, or any
 PT of its homologs or orthologs, for the treatment of e.g. deafness,
 PT osteoarthritis and abnormal cell proliferation.
 XX
 PS Disclosure; Page; 142pp; English.
 XX
 CC The invention relates to the use of atonal-associated nucleic acid or

05-DEC-2001; 2001US-00004717.
(BAYU) BAYLOR COLLEGE MEDICINE.
Zoghbi HY, Qi Y;
WPI; 2003-505253/47.
N-PSDB; AAL62224.
New compositions comprising at least one stem cell which is up- or down-regulated for expression of an atonal-associated sequence, and at least one regulatory factor, useful for treating a gastrointestinal condition, e.g. cancer.
Disclosure; Page 144-145; 157pp; English.
The invention relates to pharmaceutical compositions and methods for the therapeutic use of an atonal-associated sequence. The composition comprises of a stem cell which is up- or down-regulated for expression of an atonal-associated sequence and a regulatory factor. Atonal-associated amino acid or nucleic acid sequence is used to treat a gastrointestinal condition including cancer, damaged intestinal tissue, inflammatory bowel disease, irritable bowel syndrome, infection or necrotising enterocolitis. They are also useful for promoting mechanoreceptive cell growth, for generating hair cells, for treating hearing impairment or imbalance disorder, joint disease, abnormal proliferation (e.g. cancer), or a disease that is a result or loss of functional atonal-associated sequences. Atonal-associated sequences are also used in gene therapy. The present sequence is mouse atonal homologue 1 (Math1) protein. This sequence is used to illustrate the method of the invention
Sequence 351 AA;
XX
SQ

Query Match	87.4%; Score 1627.5; DB 6; Length 351;
Best Local Similarity	88.4%; Pred. No. 5.4e-131;
Matches 313; Conservative	11; Mismatches 27; Indels 3; Gaps 2
QY	1 MSRLHAEEWAEVKELGDHHRQOPHHLPPOPPEPPATLQAREHPVVPPELSLLDSTD 60
DB	1 MSRLHAEEWAEVKELGDHHRHHPHHV--PPLTPPPATLQARDLPVYPAELSLDSTD 58
QY	61 PRAWLAPTLQGITCTARAAQYLLHSPELGASEAAAPRDEVDGRGELVRRSSGGGASSKSPG 120
DB	59 PRAWLTPTLQGITCTARAAQYLLHSPELGASEAAAPRDEADSQGELVRR--SGCGGLSKSPG 117
QY	121 PVKVRQELCKLGGVVVDDELGCSQRAPSPSKQVNGVQKQRLLAANAERERRMEGLHNAFD 180
DB	118 PVKVRQELCKLGGVVVDDELGCSQRAPSSKQVNGVQKQRLLAANAERERRMEGLHNAFD 177
QY	181 QLBNVIPSFNNDKLSKYETLQMAQIYINALSELLQTPSGEGQPPPPASCCKDDHHLRT 240
DB	178 QLBNVIPSFNNDKLSKYETLQMAQIYINALSELLQTPNVGEQPPPTTASCCKNDHHLRT 237
QY	241 AASYEGGAGNATAAGAAQOASGSGSORPTPPGSGCRTFRFSAPASAGGYSVQLDALHFSTFDS 300
DB	238 AASYEGGAGASAVAGAQAPGCGPRPTPPGCGRTFRFSGPASSGGYSVQLDALHFPAFEDR 297
QY	301 ALTAMMAQKNLSPLPGSLIPLQVQENSKTSPRHSRDGEPSPHSHVSDSDEAS 354
DB	298 ALTAMMAQKNLSPLPGSLIPLQVQEDNSKTSPRHSRDGEPSPHSHVSDSDEAS 351

Qy	181	QLRNVIPFNNDDKGLSKYETILQMAQIYNALUSELLQTPSGGEQPPPPFASCKASDAHHLR	237
Db	178	QLRNVIPFNNDDKGLSKYETILQMAQIYNALSELLOTPNVGEOPPPTASCKNDHHLR	237
Qy	241	AASYEGGAGNATAAGCAQAAGSGSORPTPPGSCRTFRFAPASAGGYSVQLDALHFSTPDS	3000
Db	238	ASYEKGAGASAVAQAQAPGGGPRPTPPGCRTRFSGPASSGGYSVQLDALHFAPEDR	2979
Qy	301	ALTAMMAQKNLSPLPGSILOPVQEENSKTSPRHRSDGCFSPSHSHVSDSDEAS	354
Db	298	ALTAMMAQKDLSPSLPGGILOPVQEDNSKTSRHRSDGCFSPSHSHVSDSDEAS	351

RESULT 8

AAO30924	ID	AAO30924	standard; protein; 351 AA.
XX	AC	AAO30924;	
XX	DT	22-SEP-2003	(first entry)
XX	DE	Mouse atonal homologue 1 (Math1) protein #4.	

KW Atonal-associated protein; gastrointestinal; damaged intestinal tissue;
 KW cancer; inflammatory bowel disease; irritable bowel syndrome; infection;
 KW necrotising enterocolitis; hearing impairment; abnormal proliferation;
 KW joint disease; gene therapy; imbalance disorder; atonal homologue 1;
 KW Math1; mouse.

XX Mus musculus.

XX WO2003047532-A2.

XX 12-JUN-2003.

XX 03-DEC-2002; 2002WO-US041458.

XX 05-DEC-2001; 2001US-00004717.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX Zoghbi HY, Qi Y;

XX WPI; 2003-505253/47.

XX N-PSDB; AAL62204.

XX New compositions comprising at least one stem cell which is up- or down-
 regulated for expression of an atonal-associated sequence, and at least
 one regulatory factor, useful for treating a gastrointestinal condition,
 e.g. cancer.

XX Disclosure; Page 122-123; 157pp; English.

XX The invention relates to pharmaceutical compositions and methods for the
 therapeutic use of an atonal-associated sequence. The composition
 comprises of a stem cell which is up- or down-regulated for expression of
 an atonal-associated sequence and a regulatory factor. Atonal-associated
 amino acid or nucleic acid sequence is used to treat a gastrointestinal
 condition including cancer, damaged intestinal tissue, inflammatory bowel
 disease, irritable bowel syndrome, infection or necrotising enterocolitis.
 They are also useful for promoting mechanoreceptive cell growth, for
 generating hair cells, for treating hearing impairment or imbalance
 disorder, joint disease, abnormal proliferation (e.g. cancer), or a
 disease that is a result or loss of functional atonal-associated
 sequences. Atonal-associated sequences are also used in gene therapy. The
 present sequence is mouse atonal homologue 1 (Math1) protein. This
 sequence is used to illustrate the method of the invention

XX Sequence 351 AA;

Query Match 87.4%; Score 1627.5; DB 6; Length 351;
 Best Local Similarity 88.4%; Pred. No. 5.4e-131;
 Matches 313; Conservative 11; Mismatches 27; Indels 3; Gaps 2;
 QY 1 MSRLHAEWAWEKELGDRHQPQPHLPQPPPPPPATLQAREHPVVPPELSLSDTD 60
 Db 1 MSRLHAEWAWEKELGDRHQPQPHV--PPLTPPPATLQARDLPVPAELSLSDTD 58
 QY 61 PRAWLPTLQICITAAQVLLHSPELGASEAAAPRDENVGRLVRSGGGASSKSPG 120
 Db 59 PRAWLPTLQICITAAQVLLHSPELGASEAAAPRDENVGRLVRS--SGCGLSKSPG 117
 QY 121 PVKVRQLCKLGGVVVDLGCGRAPSKQVNGVKORRLAANARERRMHGLNHPD 180
 Db 118 PVKVRQLCKLGGVVVDLGCGRAPSKQVNGVKORRLAANARERRMHGLNHPD 177
 QY 181 QLRNVIPSNNDKKLSKYETLQMAQIYINALSELLQTPSGGQPPPPPPASCKSDHHRLT 240
 Db 178 QLRNVIPSNNDKKLSKYETLQMAQIYINALSELLQTPNVGQPPPTASCKNDHHRLT 237
 QY 241 AASYEGGAGNATAAGQAQSGGQRTTPGSCRTFSPASAGGYSVQLDALHFTFDS 300
 Db 238 AASYEGGAGASAVAGQAQAPGGGFRTPPGPCRTFRFGSPASSGGYSVQLDALHFPFEDR 297
 QY 301 ALTAMMAQKDLSPSLPGSILQPVQEDNSKTSRSHRSDGEFSPHSHYSDEAS 354

Db 298 ALTAMMAQKDLSPSLPGSILQPVQEDNSKTSRSHRSDGEFSPHSHYSDEAS 351

RESULT 9

AAB60366

XX AAB60366 standard; protein; 266 AA.

XX AAB60366;

XX 24-APR-2001 (first entry)

XX Zebrafish atonal homologue-1 (zath-1) protein, SEQ ID NO:40.

XX Atonal; homologue; orthologue; atonal-associated protein; deafness;
 KW hearing impairment; vestibular effect; balance disorder; osteoarthritis;
 KW cellular proliferation; cerebellar granule neuron; gene therapy;
 KW mechanoreceptive cell growth; auditory; osteopathic; cytostatic;
 KW transgenic animal.

XX Danio rerio.

XX WO2000073764-A2.

XX 07-DEC-2000.

XX 01-JUN-2000; 2000WO-US015410.

XX 01-JUN-1999; 99US-0137060P.

XX 19-JAN-2000; 2000US-0176993P.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX Zoghbi HY, Bellen H, Birmingham N, Hassan B, Ben-Arie N;

XX WPI; 2001-032190/04.

XX N-PSDB; AAF27274.

XX Therapeutic use of atonal-associated nucleic acids or amino acids, or any
 PT of its homologs or orthologs, for the treatment of e.g. deafness,
 PT osteoarthritis and abnormal cell proliferation.

XX Disclosure; Page; 142pp; English.

XX The invention relates to the use of atonal-associated nucleic acid or
 amino acid sequence, or any of its homologues or orthologues as
 therapeutic agents for the treatment of deafness, partial hearing loss,
 vestibular effects due to damage or loss of inner hair cells,
 osteoarthritis and abnormal cell proliferation. The invention also
 encompasses methods of screening for compounds which affect the
 expression of an atonal-associated nucleic acid sequence in an animal,
 and a transgenic animal in which an allele of a native atonal-associated
 gene is replaced by a heterologous nucleic acid sequence, thus
 inactivating the atonal-associated allele. The nucleic acids or proteins
 may be used in a method of treating an animal for hearing impairment,
 joint disease, balance disorders, abnormal cell proliferation, or other
 disease related to loss of a functional atonal-associated nucleic acid or
 protein. They may particularly be used to treat an animal with a
 deficiency in cerebellar granule neurons or their precursors, and may
 also be used in promoting mechanoreceptive cell growth and generating
 hair cells. The present sequence represents an atonal-associated amino
 acid sequence referred to in the invention. Note: The present sequence is
 not shown in the specification, but was obtained from GenBank

XX Sequence 266 AA;

Query Match 31.3%; Score 582; DB 4; Length 266;
 Best Local Similarity 46.5%; Pred. No. 1.8e-41;

Matches 155; Conservative 21; Mismatches 69; Indels 88; Gaps 14;

QY 12 EYKELGDHHRQPHLPQPPPPPPATLQAREHPVVPPELSLSDTDPRAWLPTLOG 71

Db 10 EVELDVQH-----SSLRGEGSKYPPALMALMASSDPRAWLAPVQAG 51

CC hair cells. The present sequence represents an atonal-associated amino
CC acid sequence referred to in the invention. Note: The present sequence is
CC not shown in the specification, but was obtained from GenBank

XX
SQ Sequence 161 AA;

Query Match 18.1%; Score 337; DB 4; Length 161;
Best Local Similarity 61.2%; Pred. No. 9.3e-21;
Matches 79; Conservative 10; Mismatches 22; Indels 18; Gaps 5;

QY 107 RRSGGAS----SKSPGVKVRQCLKLGKGVVDELGCSRQAPSSKQVGVKQRRLA 163
Db 42 RPVAGGAARVPAGAAGE---RGAAGARGGG-----GGAGPRA---QVSGVQKQRRLA 89

QY 164 ANARERRRMEGLNHAFLDQLRNVPFNNNDKLSKYETLQMAQIYINALSSELLQTPPSGGEQ 223
Db 90 ANARERRRMEGLNHAFLDQLRNVPFNNNDKLSKYETLQMAQIYISALAEHLHGPPA--- 146

QY 224 PPPPPASCK 232
Db 147 PPEPPAKAE 155

RESULT 13
AAO30946
ID AAO30946 standard; protein; 161 AA.
XX
AC AAO30946;
XX
DT 22-SEP-2003 (first entry)
XX
DE Chicken atonal homologue 1 (Cathl) protein #4.
XX
KW Atonal-associated protein; gastrointestinal; damaged intestinal tissue;
KW cancer; inflammatory bowel disease; irritable bowel syndrome; infection;
KW neurotising enterocolitis; hearing impairment; abnormal proliferation;
KW joint disease; gene therapy; imbalance disorder; atonal homologue 1;
KW Cathl; chicken.
XX
OS Gallus sp.
XX
PN WO2003047532-A2.
XX
PD 12-JUN-2003.
XX
PF 03-DEC-2002; 2002WO-US041459.
XX
PR 05-DEC-2001; 2001US-00004717.
XX
PA (BAYU) BAYLOR COLLEGE MEDICINE.
XX
PI Zoghbi HY, Qi Y;
XX
XX WPI; 2003-505253/47.
DR N-PSDB; AAJ62231.
XX
XX New compositions comprising at least one stem cell which is up- or down-
PT regulated for expression of an atonal-associated sequence, and at least
PT one regulatory factor, useful for treating a gastrointestinal condition,
PT e.g. cancer.
XX
PS Disclosure; Page 152-153; 157pp; English.

XX
CC The invention relates to pharmaceutical compositions and methods for the
CC therapeutic use of an atonal-associated sequence. The composition
CC comprises of a stem cell which is up- or down-regulated for expression of
CC an atonal-associated sequence and a regulatory factor. Atonal-associated
CC amino acid or nucleic acid sequence is used to treat a gastrointestinal
CC condition including cancer, damaged intestinal tissue, inflammatory bowel
CC disease, irritable bowel syndrome, infection or necrotising enterocolitis.
CC They are also useful for promoting mechanoreceptive cell growth, for
CC generating hair cells, for treating hearing impairment or imbalance
CC disorder, joint disease, abnormal proliferation (e.g. cancer), or a

CC disease that is a result or loss of functional atonal-associated
CC sequences. Atonal-associated sequences are also used in gene therapy. The
CC present sequence is chicken atonal homologue 1 (Cathl) protein. This
CC sequence is used to illustrate the method of the invention

XX
SQ Sequence 161 AA;

Query Match 18.1%; Score 337; DB 6; Length 161;
Best Local Similarity 61.2%; Pred. No. 9.3e-21;
Matches 79; Conservative 10; Mismatches 22; Indels 18; Gaps 5;

QY 107 RRSGGAS----SKSPGVKVRQCLKLGKGVVDELGCSRQAPSSKQVGVKQRRLA 163
Db 42 RPVAGGAARVPAGAAGE---RGAAGARGGG-----GGAGPRA---QVSGVQKQRRLA 89

QY 164 ANARERRRMEGLNHAFLDQLRNVPFNNNDKLSKYETLQMAQIYINALSSELLQTPPSGGEQ 223
Db 90 ANARERRRMEGLNHAFLDQLRNVPFNNNDKLSKYETLQMAQIYISALAEHLHGPPA--- 146

QY 224 PPPPPASCK 232
Db 147 PPEPPAKAE 155

RESULT 14
AAB60356
ID AAB60356 standard; protein; 178 AA.
XX
AC AAB60356;
XX
DT 24-APR-2001 (first entry)
XX
DE Chicken atonal homologue ngn1/ath4c protein, SEQ ID NO:19.
XX
KW Atonal; homologue; orthologue; atonal-associated protein; deafness;
KW hearing impairment; vestibular effect; balance disorder; osteoarthritis;
KW cellular proliferation; cerebellar granule neuron; gene therapy;
KW mechanoreceptive cell growth; auditory; osteopathic; cytostatic;
KW transgenic animal.
XX
OS Gallus gallus.
XX
PN WO2000073764-A2.
XX
PD 07-DEC-2000.
XX
PF 01-JUN-2000; 2000WO-US015410.
XX
PR 01-JUN-1999; 99US-0137060P.
PR 19-JAN-2000; 2000US-0176933P.
XX
PA (BAYU) BAYLOR COLLEGE MEDICINE.
XX
PI Zoghbi HY, Bellen H, Birmingham N, Hassan B, Ben-Arie N;
XX
XX WPI; 2001-032190/04.
DR N-PSDB; AAF27263.
XX
XX Therapeutic use of atonal-associated nucleic acids or amino acids, or any
PT of its homologs or orthologs, for the treatment of e.g. deafness,
PT osteoarthritis and abnormal cell proliferation.
XX
PS Disclosure; Page; 142pp; English.
XX
CC The invention relates to the use of atonal-associated nucleic acid or
CC amino acid sequence, or any of its homologues or orthologues as
CC therapeutic agents for the treatment of deafness, partial hearing loss,
CC vestibular effects due to damage or loss of inner hair cells,
CC osteoarthritis and abnormal cell proliferation. The invention also
CC encompasses methods of screening for compounds which affect the
CC expression of an atonal-associated nucleic acid sequence in an animal,
CC and a transgenic animal in which an allele of a native atonal-associated
CC gene is replaced by a heterologous nucleic acid sequence, thus

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 21:18:48 ; Search time 43.424 Seconds
(without alignments)
420.864 Million cell updates/sec

Title: US-09-980-381A-58

Perfect score: 1862
Sequence: 1 MSRLHAHAEWAWEKELGDHH.....HRSDGEPSPHSHYSDSDEAS 354

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pap.*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pap.*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pap.*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pap.*
5: /cgn2_6/ptodata/2/iaa/6C.COMB.pap.*
6: /cgn2_6/ptodata/2/iaa/6D.COMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	281	15.1	57	4	US-08-722-570-8
2	281	15.1	57	4	US-08-932-411A-8
3	217	11.7	263	4	US-08-932-411A-18
4	214	11.5	352	1	US-08-552-142A-4
5	214	11.5	352	1	US-08-910-973-4
6	214	11.5	352	4	US-09-499-227-4
7	214	11.5	352	5	PCT-US95-05741-4
8	210.5	11.3	381	1	US-08-910-973-11
9	210.5	11.3	381	4	US-09-499-227-11
10	210	11.3	382	3	US-09-234-332-9
11	209.5	11.3	237	1	US-08-910-973-13
12	209.5	11.3	237	4	US-09-499-227-13
13	206.5	11.1	379	1	US-08-552-142A-11
14	205	11.0	244	4	US-08-722-570-1
15	205	11.0	244	4	US-08-932-411A-1
16	204	11.0	380	1	US-08-552-142A-17
17	202	10.8	244	1	US-08-910-973-22
18	202	10.8	244	4	US-09-499-227-22
19	202	10.8	244	4	US-08-932-411A-14
20	202	10.8	357	1	US-08-552-142A-2
21	202	10.8	357	1	US-08-910-973-2
22	202	10.8	357	4	US-09-499-227-2
23	202	10.8	357	5	PCT-US95-05741-2
24	201.5	10.8	214	4	US-08-932-411A-20
25	201.5	10.8	382	1	US-08-910-973-17
26	201.5	10.8	382	4	US-09-499-227-17
27	200	10.7	356	1	US-08-552-142A-15

28	200	10.7	356	1	US-08-910-973-15	Sequence 15, Appl
29	200	10.7	356	4	US-09-499-227-15	Sequence 15, Appl
30	199	10.7	356	3	US-09-234-332-7	Sequence 7, Appl
31	196.5	10.6	214	4	US-08-722-570-2	Sequence 2, Appl
32	196.5	10.6	214	4	US-08-932-411A-2	Sequence 2, Appl
33	196	10.5	57	4	US-08-722-570-9	Sequence 9, Appl
34	194	10.5	57	4	US-08-932-411A-9	Sequence 9, Appl
35	194	10.4	356	3	US-09-234-332-8	Sequence 8, Appl
36	182	9.8	156	1	US-08-552-142A-9	Sequence 9, Appl
37	182	9.8	156	1	US-08-910-973-9	Sequence 9, Appl
38	182	9.8	156	4	US-09-499-227-9	Sequence 9, Appl
39	182	9.8	156	5	PCT-US95-05741-9	Sequence 9, Appl
40	173	9.3	57	4	US-08-722-570-5	Sequence 5, Appl
41	173	9.3	57	4	US-08-932-411A-5	Sequence 5, Appl
42	169	9.1	57	4	US-08-722-570-6	Sequence 6, Appl
43	169	9.1	57	4	US-08-932-411A-6	Sequence 6, Appl
44	166	8.9	161	5	PCT-US95-05741-11	Sequence 11, Appl
45	162	8.7	57	4	US-08-722-570-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-08-722-570-8
; Sequence 8, Application US/08722570
; Patent No. 6555337
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Ma, Qifu
; TITLE OF INVENTION: NEUROGENIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/722,570
; FILING DATE: 27-SEP-1996
; CLASSIFICATION: 5365
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-63902/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-722-570-8

Query Match 15.1%; Score 281; DB 4; Length 57;
Best Local Similarity 96.5%; Pred. No. 2.5e-19;
Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 158 KORRLAANARERRRMHGLNHFQDLRNVPSPNNDKLSKYETLQMAQIYINALSEL 214
DB 1 KNRRLAANARERRRMHGLNHFQDLRNVPSPNNDKLSKYETLQMAQIYINALSEI 57

RESULT 2

US-08-932-411A-8
; Sequence 8, Application US/08932411A
; Patent No. 6566496
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Ma, Qifu
; TITLE OF INVENTION: NEUROGENIN
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr Hobbach Test Albritton & Herbert LLP
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/932,411A
; FILING DATE: 15-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/772,009
; FILING DATE: 19-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/722,570
; FILING DATE: 19-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-63902-3/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-932-411A-18

Query Match 15.1%; Score 281; DB 4; Length 57;
Best Local Similarity 96.5%; Pred. No. 2.5e-19;
Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 158 KORLLAANARERRRMHGLNHFADQLRNVPSPNNKKLSKYETLQMAQIYNALSEL 214
Db 1 KNRLAANARERRRMHGLNHFADQLRNVPSPNNKKLSKYETLQMAQIYNALSEL 57

RESULT 3

US-08-932-411A-18
; Sequence 18, Application US/08932411A
; Patent No. 6566496
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Ma, Qifu
; TITLE OF INVENTION: NEUROGENIN
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr Hobbach Test Albritton & Herbert LLP
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States

ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/932,411A
; FILING DATE: 15-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/772,009
; FILING DATE: 19-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/722,570
; FILING DATE: 19-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-63902-3/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-932-411A-18

Query Match 11.7%; Score 217; DB 4; Length 263;
Best Local Similarity 28.7%; Pred. No. 2.4e-12;
Matches 98; Conservative 25; Mismatches 89; Indels 130; Gaps 15;

QY 40 TLQAREHPVYPPELSLLDSTDP-RAWLAP-----TLOGICTARAAOYL 81
Db 7 TLELKEE---EEVLMLLGSASASATLTWSSSSADEEDELRRPSSARGQGAEEQGV 63
QY 82 LHSPELGASAAAPRDVGRGELVRRSSGSSSSKSPGVKVRQOLCKLKGVVVDLGL 141
Db 64 QGSPASGA-----GGC-----RPGRLILGLMHECKRRP----- 90
QY 142 CSRQRAPS-----SKOVNGVQKORLLAANARERRRMHGLNHFADQLRNVPSPNNKKLS 196
Db 91 -SRRAVSARGAKTAETVQRIKTRRLKANRRERNRMHNLNAALDALREVLPFPEDAKLT 149
QY 197 KYETLQMAQIYINALSELQTPSGGEGEQPPPPPPASCKSDHHLRLTAASYEGGAGNATAA-- 254
Db 150 KIETLRFHNYIWTETLR-----LADHCAGAGGLOGALF 185
QY 255 -----GAQASGGSORPTTPGSCRTFRFASAPASAGYSVQLDALHFTTFEDSALTAM 305
Db 186 TEAVLLSPGAALGASGDSFPPSPWSCT---NSPASSN-----STGPYSC----- 228
QY 306 MAQKNLSPSLPGSII---QPVQENSKTSPSRHSRSDGFSPH 344
Db 229 ----TLSPASPGSDVDYQWP-----PPPEKHR-----YAPH 255

RESULT 4

US-08-552-142A-4
; Sequence 4, Application US/08552142A
; Patent No. 5695995
; GENERAL INFORMATION:
; APPLICANT: Weintraub, Harold M.
; APPLICANT: Lee, Jacqueline E.
; APPLICANT: Tapscott, Stephen J.
; APPLICANT: Hollenberg, Stanley M.
; TITLE OF INVENTION: Neurogenic Differentiation (NeuroD) Genes
; TITLE OF INVENTION: and Proteins
; NUMBER OF SEQUENCES: 20

Wed Sep 22 12:21:30 2004

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/499,227
 FILING DATE: 05-August-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/239,238
 FILING DATE: 06-May-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/US95/05741
 FILING DATE: 08-May-1995
 APPLICATION NUMBER: PCT/US96/17532
 FILING DATE: 30-October-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/910,973
 FILING DATE: 07-August-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Sheiness, Diana K.
 REGISTRATION NUMBER: 35,356
 REFERENCE/DOCKET NUMBER: PHCR-1-12742
 TELEPHONE: 206-682-8100; 206-224-0735 (direct)
 TELEFAX: 206-225-0779
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 381 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 US-09-499-227-11

Query Match 11.3%; Score 210.5; DB 4; Length 381;
 Best Local Similarity 26.4%; Pred. No. 1.7e-11;
 Matches 98; Conservative 42; Mismatches 126; Indels 105; Gaps 17;

QY	22	QPOPHLPQPPPPPPPPATLQAREHVPYPPPELSLLDTPRAWLAPTLOGICTARAAQYL	81
DB	28	EPKSDKGDAPPPPPAP-----GPGAPG--PARAAKPV	58
QY	82	LHSPELGASAAAPRDEVDGRELVRSSGASSKSPGVKVRQOLCKLGGVVVDDEL	141
DB	59	---PLRGEETATLAEVKEGEL-----GGEDEE-----E	99
QY	142	CSRQAPSSKQVNGVQ---KORLAANARRRRMHGLNHAFLDQLRNVIPIFNNDKLSK	197
DB	100	RPKRGPKKRMKTKARLERSKLKQKANARERNRMDLNAALDNLKRVKVPVCSYKTKLSK	159
QY	198	YETLQMAQIYNALSEL-----QTPSGGEQPPPP--PASC-KSDHHLRTAA	242
DB	160	IETLRLAKNYIWLSEILRSCKRDLVSVYVTLCKGLSQPTNLVAGCLNSRNFLTEQ	219
QY	243	SYEG-----GAGNATA-----AGAQ-QASGG-----SORPTPGSCRTFSAPA	280
DB	220	GADGAGRPHGSGGPFAMHPYPYPCSLAGACQAAAGLGGGAHAALRTHGYCAAYETLYA	279
QY	281	SAGYSVQLDALHFTFEDSALTAMMAQNLSPSLPGSILQPVQENSKTSRSHRSDGE	340
DB	280	AAGGGASPD--YNSSEYEGPLSPPLCLNGNFS-----LKQDSSPDHEKSVHYSMH	328
QY	341	FSP-----HSH 346	
DB	329	YSAALPGSRHG 339	

RESULT 10
 US-09-234-332-9
 ; Sequence 9, Application US/09234332A
 ; Patent No. 6087168
 ; GENERAL INFORMATION:

APPLICATION NUMBER: WO PCT/US95/05741
 FILING DATE: 08-MAY-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US96/17532
 FILING DATE: 30-October-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Sheiness, Diana K.
 REGISTRATION NUMBER: 35,356
 REFERENCE/DOCKET NUMBER: PHCR-1-10958
 TELEPHONE: 206-682-8100; 206-224-0735 (direct)
 TELEFAX: 206-225-0779
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 381 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 US-08-910-973-11

Query Match 11.3%; Score 210.5; DB 1; Length 381;
 Best Local Similarity 26.4%; Pred. No. 1.7e-11;
 Matches 98; Conservative 42; Mismatches 126; Indels 105; Gaps 17;

QY	22	QPOPHLPQPPPPPPPPATLQAREHVPYPPPELSLLDTPRAWLAPTLOGICTARAAQYL	81
DB	28	EPKSDKGDAPPPPPAP-----GPGAPG--PARAAKPV	58
QY	82	LHSPELGASAAAPRDEVDGRELVRSSGASSKSPGVKVRQOLCKLGGVVVDDEL	141
DB	59	---PLRGEETATLAEVKEGEL-----GGEDEE-----E	99
QY	142	CSRQAPSSKQVNGVQ---KORLAANARRRRMHGLNHAFLDQLRNVIPIFNNDKLSK	197
DB	100	RPKRGPKKRMKTKARLERSKLKQKANARERNRMDLNAALDNLKRVKVPVCSYKTKLSK	159
QY	198	YETLQMAQIYNALSEL-----QTPSGGEQPPPP--PASC-KSDHHLRTAA	242
DB	160	IETLRLAKNYIWLSEILRSCKRDLVSVYVTLCKGLSQPTNLVAGCLNSRNFLTEQ	219
QY	243	SYEG-----GAGNATA-----AGAQ-QASGG-----SORPTPGSCRTFSAPA	280
DB	220	GADGAGRPHGSGGPFAMHPYPYPCSLAGACQAAAGLGGGAHAALRTHGYCAAYETLYA	279
QY	281	SAGYSVQLDALHFTFEDSALTAMMAQNLSPSLPGSILQPVQENSKTSRSHRSDGE	340
DB	280	AAGGGASPD--YNSSEYEGPLSPPLCLNGNFS-----LKQDSSPDHEKSVHYSMH	328
QY	341	FSP-----HSH 346	
DB	329	YSAALPGSRHG 339	

RESULT 9
 US-09-499-227-11
 ; Sequence 11, Application US/09499227
 ; Patent No. 644463
 ; GENERAL INFORMATION:
 ; APPLICANT: Tapscott, Stephen J.
 ; APPLICANT: Olson, James M.
 ; TITLE OF INVENTION: Expression of Neurogenic bHLH Genes in Primitive Neuroectoder
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Christensen O'Connor Johnson Kindness PLLC
 ; STREET: 1420 Fifth Avenue, Suite 2800
 ; CITY: Seattle
 ; STATE: WA
 ; COUNTRY: USA
 ; ZIP: 98101-2347
 ; COMPUTER READABLE FORM: disk
 ; MEDIUM TYPE: Floppy disk

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/910,973
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/239,238
; FILING DATE: 06-MAY-1994
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: WO PCT/US95/05741
; FILING DATE: 08-MAY-1995
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: PCT/US96/17532
; FILING DATE: 30-October-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sheiness, Diana K.
; REGISTRATION NUMBER: 35,356
; REFERENCE/DOCKET NUMBER: PHCR-1-10958
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-682-8100; 206-224-0735 (direct)
; TELEFAX: 206-225-0779
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 237 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-910-973-13

Query Match          11.3%; Score 209.5; DB 1;
Best Local Similarity 33.8%; Pred No. 1.1e-11;
Matches      70; Conservative    26; Mismatches    64;

Qy 107 RRSGGAS--SSKSPGVKVREOLCKLKGVVVDGLGCSRQRQ-----RTR
Db 49 RGAPNISRASEVPGAQQDEQERRRRRG-----AESWGSGAAA
Qy 166 ARERRMHGLNHAFDLNRNVIPESFNNDKKLSKYETLOWAQD
Db 99 DRERNRHMLNALDALSRSLVLPSPDDTKLTIKITLRFAYN
Qy 221 GEQ---PP-----PPPASCSDHHHLRTAASEYEGAGNA
Db 159 GARERLLPQCVPCLPGPPPSAD-----AESWGSGAAA
Qy 267 TPGSCRTRFSPASAGYSVOLDAIH 293
Db 213 GDP-----VFSTP-----SLPKDLLH 228

RESULT 12
US-09-499-227-13
; Sequence 13, Application US/09499227
; Patent No. 644463
; GENERAL INFORMATION:
; APPLICANT: Tapscott, Stephen J.
; APPLICANT: Olson, James M.
; TITLE OF INVENTION: Expression of Neurogenic bHLH
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen O'Connor Johnson Kindness
; STREET: 1420 Fifth Avenue, Suite 2800
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101-2347
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/552,142A
FILING DATE: 02-NOV-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,238
FILING DATE: 06-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US95/05741
FILING DATE: 08-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Broderick, Thomas F.
REGISTRATION NUMBER: 31,332
REFERENCE/DOCKET NUMBER: FHC-1-8933
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-682-8100
TELEFAX: 206-225-0709
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 379 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-552-142A-11

Query Match 11.1%; Score 206.5; DB 1; Length 379;
Best Local Similarity 25.5%; Pred. No. 4e-11;
Matches 94; Conservative 38; Mismatches 134; Indels 103; Gaps 14;

QY 22 QPQHLLPQPPPPATLQAREHPVYPPELSLIDTDPRAWLPTLQGGICTARAAQYL 81
DB 28 BPRSDKGAPPPPPAP-----GPGAPG--PARAAKPV 58
QY 82 LHSPLGSAEAAAPRDEVDGELVRRSSGGASSKSPGVKVRQLCKLKGVVVDLGL 141
DB 59 ---PLRGEETATLAEVKEGEL---GGEDEEE---EEERGLDEAG-----E 99
QY 142 CSRQAPSKQVNGVQ---KORLAANARERRRMHGLNHAEDQLRNVIPFNNDKLSK 197
DB 100 RPKRGPKRKMTKARLERSKLROKANARERNRMDLNAALDNLRKVVPCYKTKLSK 159
QY 198 YETLQMAQIYNALSELLA-----QTPSGEQPPPP--PASC-KSDHHHLRTAA 242
DB 160 IETLRKNIWALSEILRSKGRPDVSVYQTLCKGLSQPTTNLVAGCLQLNSRNFLEQ 219
QY 243 SYEGGA-----GNATAAGAAQAGSGSORPTPPGSCRTFRFAPASA 282
DB 220 GRDAXRPHSGGPPFAMHPYPYPCSRGGRTVFGAAAWAAGARLRTGCAAYETLYAAA 279
QY 283 GGYSVOLDALHFTFSDSALTAMMAQKNLSPSLPGSILQPVQEENSKTSPRS-----HRS 337
DB 280 GGGGASPD--YNSSEYEGPLSPCLCLNGFS-----LKQDSSPDHEKSHYHSMHYS 328
QY 338 DGEFSPHSH 346
DB 329 GCPGSRHGH 337

RESULT 14

US-08-722-570-1
Sequence 1, Application US/08722570
Patent No. 6555337
GENERAL INFORMATION:
APPLICANT: Anderson, David J.
APPLICANT: Ma, Qifu
TITLE OF INVENTION: NEUROGENIN
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400

APPLICATION NUMBER: US/09/499,227
FILING DATE: 05-August-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,238
FILING DATE: 06-May-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US95/05741
FILING DATE: 08-May-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/17532
FILING DATE: 30-October-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/910,973
FILING DATE: 07-August-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sheiness, Diana K.
REGISTRATION NUMBER: 35,356
REFERENCE/DOCKET NUMBER: FHC-1-12742
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-682-8100; 206-224-0735 (direct)
TELEFAX: 206-225-0779
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 237 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-499-227-13

Query Match 11.3%; Score 209.5; DB 4; Length 237;
Best Local Similarity 33.8%; Pred. No. 1.1e-11;
Matches 70; Conservative 26; Mismatches 64; Indels 47; Gaps 9;

QY 107 RASGGAS-SKSGPGVKVRQOLCKLKGVVVDLGCGRAPSKQVNGVQKORRLAAN 165
DB 49 RGAENIASEVPAQDEQERRRG-----RTRVRSALLHLRSRRVKAN 98
QY 166 ARERRMHGLNHAFLQRLNVPSPNNDKLSKYETLQMAQIYNALSELLQ-----TPSG 220
DB 99 DERNRMNNAALDARSVLSPFPDDTKLTETLRFAINYIWALETLRADQGLPGG 158
QY 221 GQV-----PPASCKSDHHHLRTAASVGGAGNATAAGAAQAGSGSQ-----RP 266
DB 159 GARERLLPQVCPFLGPPSPASD-----AESWGGAAASPLDPPSPAASEFTYRP 212
QY 267 TPGSCRTFRFAPASAGGYSVOLDALH 293
DB 213 GDP-----VFSFP-----SLPKDLH 228

RESULT 13

US-08-552-142A-11
Sequence 11, Application US/08552142A
Patent No. 5695995
GENERAL INFORMATION:
APPLICANT: Weintraub, Harold M.
APPLICANT: Lee, Jacqueline E.
APPLICANT: Tapscott, Stephen J.
APPLICANT: Hollenberg, Stanley M.
TITLE OF INVENTION: Neurogenic Differentiation (NeuroD) Genes
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Christensen O'Connor Johnson KindnessPLLC
STREET: 1420 Fifth Avenue, Suite 2800
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101-2347
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/722,570
; FILING DATE: 27-SEP-1996
; CLASSIFICATION: 5365
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-63902/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 244 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-722-570-1

Query Match      11.0%; Score 205; DB 4; Length 244;
Best Local Similarity 34.2%; Pred. No. 2.9e-11;
Matches 66; Conservative 24; Mismatches 69; Indels 34; Gaps 6;

QY 95 PRDEVDRGELVRRSSGGAS-SSKSPGPVKVREQLCKLGGVVVDELGCSSRQAPSSKQV 153
Db 38 PLASTSGLSVPARRSAPTLSGASNVGGQDEEGERRRRG-----RARVRSEALL 87
QY 154 NGVQKORRLAANARERRRMHGLNHFQDLNRNIPSFNNDKKLSKYETLQMAQIYINALSE 213
Db 88 HSLRRSRVKANDRRNRNMHNLNAALDALRSVLPSPDDTKLTETLRFAYNIWALAE 147
QY 214 LLQ-----TPSGEQ-----PP-----PPASCKSDHHHLRTAASVEGGAGNATAAGAOQ 258
Db 148 TURLADQGLPGGARERLLPQCVCPLPGPPSPASD-----TESWGSAGAAASPCATV 199
QY 259 ASGGSQRPTPPGS 271
Db 200 ASPLSDPPSPSAS 212

```

RESULT 15

```

US-08-932-411A-1
; Sequence 1, Application US/08932411A
; Patent No. 6566496
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Ma, Qifu
; TITLE OF INVENTION: NEUROGENIN
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr Hobbach Test Albritton & Herbert LLP
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/932,411A
; FILING DATE: 15-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/772,009
; FILING DATE: 19-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/722,570
; FILING DATE: 19-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-63902-3/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 244 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-932-411A-1

```

```

Query Match      11.0%; Score 205; DB 4; Length 244;
Best Local Similarity 34.2%; Pred. No. 2.9e-11;
Matches 66; Conservative 24; Mismatches 69; Indels 34; Gaps 6;

QY 95 PRDEVDRGELVRRSSGGAS-SSKSPGPVKVREQLCKLGGVVVDELGCSSRQAPSSKQV 153
Db 38 PLASTSGLSVPARRSAPTLSGASNVGGQDEEGERRRRG-----RARVRSEALL 87
QY 154 NGVQKORRLAANARERRRMHGLNHFQDLNRNIPSFNNDKKLSKYETLQMAQIYINALSE 213
Db 88 HSLRRSRVKANDRRNRNMHNLNAALDALRSVLPSPDDTKLTETLRFAYNIWALAE 147
QY 214 LLQ-----TPSGEQ-----PP-----PPASCKSDHHHLRTAASVEGGAGNATAAGAOQ 258
Db 148 TURLADQGLPGGARERLLPQCVCPLPGPPSPASD-----TESWGSAGAAASPCATV 199
QY 259 ASGGSQRPTPPGS 271
Db 200 ASPLSDPPSPSAS 212

```

Search completed: September 21, 2004, 21:25:51
Job time : 44.424 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 21:24:21 ; Search time 161.424 Seconds
(without alignments)
704.240 Million cell updates/sec

Title: US-09-980-381A-58

Perfect score: 1862
Sequence: 1 MSRLHAEAEVKEVKGDLHH.....HRSDGEFPHSHYSDEAS 354

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1342398 seqs, 321133274 residues

Total number of hits satisfying chosen parameters: 1342398

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/prodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/prodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/prodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/prodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/prodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/prodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/prodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1862	100.0	354	13	US-10-004-717-2
2	1862	100.0	354	13	US-10-004-717-58
3	1862	100.0	354	16	US-10-373-249-3
4	1862	100.0	354	16	US-10-373-249-4
5	1627.5	87.4	351	13	US-10-004-717-11
6	1627.5	87.4	351	13	US-10-004-717-46
7	1627.5	87.4	351	16	US-10-373-249-1
8	1627.5	87.4	351	16	US-10-373-249-2
9	582	31.3	266	13	US-10-004-717-40
10	537	28.8	259	13	US-10-004-717-66
11	337	18.1	161	13	US-10-004-717-60
12	281	15.1	57	8	US-08-722-570-8
13	281	15.1	57	12	US-10-425-259-8
14	228.5	12.3	178	13	US-10-004-717-19
15	224.5	12.1	213	13	US-10-004-717-21

16	222	11.9	250	16	US-10-642-093-25	Sequence 25, Appl
17	222	11.9	250	16	US-10-795-002-25	Sequence 25, Appl
18	219	11.8	208	12	US-10-403-571-12	Sequence 12, Appl
19	218	11.7	263	13	US-10-004-717-7	Sequence 7, Appl
20	218	11.7	263	13	US-10-004-717-38	Sequence 38, Appl
21	217	11.7	151	13	US-10-004-717-36	Sequence 36, Appl
22	217	11.7	263	13	US-10-425-259-18	Sequence 18, Appl
23	217	11.7	263	13	US-10-004-717-31	Sequence 31, Appl
24	214	11.5	352	16	US-10-654-102-6	Sequence 6, Appl
25	214	11.5	352	16	US-10-654-102-16	Sequence 16, Appl
26	214	11.5	352	16	US-10-654-102-27	Sequence 27, Appl
27	214	11.5	352	16	US-10-654-102-38	Sequence 38, Appl
28	214	11.5	352	16	US-10-654-102-43	Sequence 43, Appl
29	213.5	11.5	237	16	US-10-642-093-23	Sequence 23, Appl
30	213.5	11.5	237	16	US-10-642-102-42	Sequence 42, Appl
31	213.5	11.5	237	16	US-10-795-002-23	Sequence 23, Appl
32	213	11.4	312	13	US-10-004-717-64	Sequence 64, Appl
33	212.5	11.4	189	13	US-10-004-717-17	Sequence 17, Appl
34	211.5	11.4	325	15	US-10-120-801-96	Sequence 96, Appl
35	210.5	11.3	381	15	US-10-435-696-32	Sequence 32, Appl
36	210.5	11.3	381	16	US-10-654-102-19	Sequence 19, Appl
37	210	11.3	350	16	US-10-654-102-11	Sequence 11, Appl
38	210	11.3	350	16	US-10-654-102-12	Sequence 12, Appl
39	210	11.3	350	16	US-10-654-102-13	Sequence 13, Appl
40	210	11.3	382	16	US-10-654-102-31	Sequence 31, Appl
41	209.5	11.3	237	16	US-10-642-093-21	Sequence 21, Appl
42	209.5	11.3	237	16	US-10-795-002-21	Sequence 21, Appl
43	207.5	11.1	355	16	US-10-654-102-9	Sequence 9, Appl
44	207	11.1	138	13	US-10-004-717-50	Sequence 50, Appl
45	206.5	11.1	379	16	US-10-654-102-36	Sequence 36, Appl

ALIGNMENTS

RESULT 1

US-10-004-717-2

; Sequence 2, Application US/10004717
; Publication No. US20020192665A1

; GENERAL INFORMATION:

; APPLICANT: ZOOHBI, HUDA Y.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC USE OF AN

; TITLE OF INVENTION: ATONAL ASSOCIATED SEQUENCE FOR DEAFNESS,
OSTEOARTHRITIS, AND ABNORMAL CELL PROLIFERATION

; FILE REFERENCE: P01899US4

; CURRENT APPLICATION NUMBER: US/10/004,717

; CURRENT FILING DATE: 2002-08-16

; PRIOR APPLICATION NUMBER: 09/585,645

; PRIOR FILING DATE: 2000-06-01

; PRIOR APPLICATION NUMBER: 60/176,993

; PRIOR FILING DATE: 2000-01-19

; PRIOR APPLICATION NUMBER: 60/137,060

; PRIOR FILING DATE: 1999-06-01

; NUMBER OF SEQ ID NOS: 69

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 354

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-004-717-2

Query Match 100.0%; Score 1862; DB 13; Length 354;

Best Local Similarity 100.0%; Pred. No. 1.7e-135;

Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSRLHAEAEVKEVKGDLHHRQPHLPQPPPPQPPATLQAREHPVYPPELSLLDSTD 60

Db 1 MSRLHAEAEVKEVKGDLHHRQPHLPQPPPPQPPATLQAREHPVYPPELSLLDSTD 60

Qy 61 PRAWLAPTLOGICTARAAQYLLHSPELGASAAAPRDVGRLVRRSSGGSSSSKSPG 120

Db 61 PRAWLAPTLOGICTARAAQYLLHSPELGASAAAPRDVGRLVRRSSGGSSSSKSPG 120

QY 121 PVKREQLCKLGGVVVDELGCSRORAPSSKQVNGVQKORRLAANARERRRMHGLNHAFD 180
Db 121 PVKREQLCKLGGVVVDELGCSRORAPSSKQVNGVQKORRLAANARERRRMHGLNHAFD 180
QY 181 QLRNVIPSFNNKKLSKYETLQMAQIYNALSELLQTPSGSGEQPPPPPPASCKSDHHHLRT 240
Db 181 QLRNVIPSFNNKKLSKYETLQMAQIYNALSELLQTPSGSGEQPPPPPPASCKSDHHHLRT 240
QY 241 AASYEGGAGNATAAGAAQASGGSORPTPGSCRTFRSAPASAGYSVQLDALHFTFEDS 300
Db 241 AASYEGGAGNATAAGAAQASGGSORPTPGSCRTFRSAPASAGYSVQLDALHFTFEDS 300
QY 301 ALTAMMAQKNLSPLPGSILQPVQENSKTSPRSHRSDGEFSPHSHYSDSDEAS 354
Db 301 ALTAMMAQKNLSPLPGSILQPVQENSKTSPRSHRSDGEFSPHSHYSDSDEAS 354

RESULT 2

US-10-004-717-58
; Sequence 58, Application US/10004717
; Publication No. US20020192665A1
; GENERAL INFORMATION:
; APPLICANT: ZOGHEI, HUDA Y.
; APPLICANT: YANG, QI
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC USE OF AN
; TITLE OF INVENTION: ATONAL ASSOCIATED SEQUENCE FOR DEAFNESS.
; TITLE OF INVENTION: OSTEOARTHRITIS, AND ABNORMAL CELL PROLIFERATION
; FILE REFERENCE: P01895US4
; CURRENT APPLICATION NUMBER: US/10/004,717
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 09/585,645
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: 60/176,993
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: 60/137,060
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 58
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-004-717-58

Query Match 100.0%; Score 1862; DB 13; Length 354;
Best Local Similarity 100.0%; Pred. No. 1.7e-135;
Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSRLHAEEWAEEVKELGDHHRQPOPHLPQPPPPPPATLQAREHPVYPPELSLLDSTD 60
Db 1 MSRLHAEEWAEEVKELGDHHRQPOPHLPQPPPPPPATLQAREHPVYPPELSLLDSTD 60
QY 61 PRAWLAPTLOGICTARAAQYLLHSPELGASAAAAPRDEVDGRLVRRSSGGASSKSPG 120
Db 61 PRAWLAPTLOGICTARAAQYLLHSPELGASAAAAPRDEVDGRLVRRSSGGASSKSPG 120
QY 121 PVKREQLCKLGGVVVDELGCSRORAPSSKQVNGVQKORRLAANARERRRMHGLNHAFD 180
Db 121 PVKREQLCKLGGVVVDELGCSRORAPSSKQVNGVQKORRLAANARERRRMHGLNHAFD 180
QY 181 QLRNVIPSFNNKKLSKYETLQMAQIYNALSELLQTPSGSGEQPPPPPPASCKSDHHHLRT 240
Db 181 QLRNVIPSFNNKKLSKYETLQMAQIYNALSELLQTPSGSGEQPPPPPPASCKSDHHHLRT 240
QY 241 AASYEGGAGNATAAGAAQASGGSORPTPGSCRTFRSAPASAGYSVQLDALHFTFEDS 300
Db 241 AASYEGGAGNATAAGAAQASGGSORPTPGSCRTFRSAPASAGYSVQLDALHFTFEDS 300
QY 301 ALTAMMAQKNLSPLPGSILQPVQENSKTSPRSHRSDGEFSPHSHYSDSDEAS 354
Db 301 ALTAMMAQKNLSPLPGSILQPVQENSKTSPRSHRSDGEFSPHSHYSDSDEAS 354

RESULT 3

US-10-373-249-3
; Sequence 3, Application US/10373249
; Publication No. US20040166091A1
; GENERAL INFORMATION:
; APPLICANT: BROUGH, DOUGLAS E
; TITLE OF INVENTION: MATERIALS AND METHODS FOR TREATING DISORDERS OF THE EAR
; FILE REFERENCE: 214680
; CURRENT APPLICATION NUMBER: US/10/373,249
; CURRENT FILING DATE: 2003-02-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-373-249-3

Query Match 100.0%; Score 1862; DB 16; Length 354;
Best Local Similarity 100.0%; Pred. No. 1.7e-135;
Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSRLHAEEWAEEVKELGDHHRQPOPHLPQPPPPPPATLQAREHPVYPPELSLLDSTD 60
Db 1 MSRLHAEEWAEEVKELGDHHRQPOPHLPQPPPPPPATLQAREHPVYPPELSLLDSTD 60
QY 61 PRAWLAPTLOGICTARAAQYLLHSPELGASAAAAPRDEVDGRLVRRSSGGASSKSPG 120
Db 61 PRAWLAPTLOGICTARAAQYLLHSPELGASAAAAPRDEVDGRLVRRSSGGASSKSPG 120
QY 121 PVKREQLCKLGGVVVDELGCSRORAPSSKQVNGVQKORRLAANARERRRMHGLNHAFD 180
Db 121 PVKREQLCKLGGVVVDELGCSRORAPSSKQVNGVQKORRLAANARERRRMHGLNHAFD 180
QY 181 QLRNVIPSFNNKKLSKYETLQMAQIYNALSELLQTPSGSGEQPPPPPPASCKSDHHHLRT 240
Db 181 QLRNVIPSFNNKKLSKYETLQMAQIYNALSELLQTPSGSGEQPPPPPPASCKSDHHHLRT 240
QY 241 AASYEGGAGNATAAGAAQASGGSORPTPGSCRTFRSAPASAGYSVQLDALHFTFEDS 300
Db 241 AASYEGGAGNATAAGAAQASGGSORPTPGSCRTFRSAPASAGYSVQLDALHFTFEDS 300
QY 301 ALTAMMAQKNLSPLPGSILQPVQENSKTSPRSHRSDGEFSPHSHYSDSDEAS 354
Db 301 ALTAMMAQKNLSPLPGSILQPVQENSKTSPRSHRSDGEFSPHSHYSDSDEAS 354

RESULT 4

US-10-373-249-4
; Sequence 4, Application US/10373249
; Publication No. US20040166091A1
; GENERAL INFORMATION:
; APPLICANT: BROUGH, DOUGLAS E
; TITLE OF INVENTION: MATERIALS AND METHODS FOR TREATING DISORDERS OF THE EAR
; FILE REFERENCE: 214680
; CURRENT APPLICATION NUMBER: US/10/373,249
; CURRENT FILING DATE: 2003-02-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-373-249-4

Query Match 100.0%; Score 1862; DB 16; Length 354;
Best Local Similarity 100.0%; Pred. No. 1.7e-135;
Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSRLHAEEWAEEVKELGDHHRQPOPHLPQPPPPPPATLQAREHPVYPPELSLLDSTD 60
Db 1 MSRLHAEEWAEEVKELGDHHRQPOPHLPQPPPPPPATLQAREHPVYPPELSLLDSTD 60

QY 61 PRAWLAPTLOGICTARAAQYLLHSPGELGASAAAAPRDEVDGREGELVRRSSGASSKSPG 120
Db 61 PRAWLAPTLOGICTARAAQYLLHSPGELGASAAAAPRDEVDGREGELVRRSSGASSKSPG 120
QY 121 PVKREOLCKLGGVVDELGCSRQAPSSKQVNGVQKORRLAANARERRRMHGLNHAFF 180
Db 121 PVKREOLCKLGGVVDELGCSRQAPSSKQVNGVQKORRLAANARERRRMHGLNHAFF 180
QY 181 QLRNVIPSFNNDKKLSKYETLQMAQIYINALSELLQTPSGGEGQPPPPASCKSDHHHLRT 240
Db 181 QLRNVIPSFNNDKKLSKYETLQMAQIYINALSELLQTPSGGEGQPPPPASCKSDHHHLRT 240
QY 241 AASYEGGAGNATAAGAAQAGSGSQRTPPGSCRTFRSAPASAGGYSVOLDALHFFSTFEDS 300
Db 241 AASYEGGAGNATAAGAAQAGSGSQRTPPGSCRTFRSAPASAGGYSVOLDALHFFSTFEDS 300
QY 301 ALTAMMAOKNLSPLPGSLQPVQENSKTSRSHRSDGFEFSPHSHYSDEAS 354
Db 301 ALTAMMAOKNLSPLPGSLQPVQENSKTSRSHRSDGFEFSPHSHYSDEAS 354

RESULT 5

US-10-004-717-11

; Sequence 11, Application US/10004717
; Publication No. US20020192665A1
; GENERAL INFORMATION:
; APPLICANT: ZOGHEI, HUDA Y.

; APPLICANT: YANG, QI

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC USE OF AN

; TITLE OF INVENTION: ATONAL ASSOCIATED SEQUENCE FOR DEAFNESS,

; FILE REFERENCE: P01899054

; CURRENT FILING DATE: 2002-08-16

; PRIOR APPLICATION NUMBER: US/10/004,717

; PRIOR FILING DATE: 2000-06-01

; PRIOR APPLICATION NUMBER: 60/176,993

; PRIOR FILING DATE: 2000-01-19

; PRIOR APPLICATION NUMBER: 60/137,060

; PRIOR FILING DATE: 1999-06-01

; NUMBER OF SEQ ID NOS: 69

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 11

; LENGTH: 351

; TYPE: PRT

; ORGANISM: Mus musculus

US-10-004-717-11

Query Match

Best Local Similarity 87.4%; Score 1627.5; DB 13; Length 351;

Matches 313; Conservative 11; Mismatches 27; Indels 3; Gaps 2;

QY 1 MSRLHAEWEAEVKELGDHHRQPHHLPPPPPPPPATLQAREHPVYPPELSLLDSTD 60

Db 1 MSRLHAEWEAEVKELGDHHRQPHHV--PPLTPPPATLQARDLPVYPPELSLLDSTD 58

QY 61 PRAWLAPTLOGICTARAAQYLLHSPGELGASAAAAPRDEVDGREGELVRRSSGASSKSPG 120

Db 59 PRAWLPTTLOGICTARAAQYLLHSPGELGASAAAAPRDEADSQGELVRR--SGCGGLSKSPG 117

QY 121 PVKREOLCKLGGVVDELGCSRQAPSSKQVNGVQKORRLAANARERRRMHGLNHAFF 180

Db 118 PVKREOLCKLGGVVDELGCSRQAPSSKQVNGVQKORRLAANARERRRMHGLNHAFF 177

QY 181 QLRNVIPSFNNDKKLSKYETLQMAQIYINALSELLQTPSGGEGQPPPPASCKSDHHHLRT 240

Db 178 QLRNVIPSFNNDKKLSKYETLQMAQIYINALSELLQTPNVGEGQPPPTASCKNDHHHLRT 237

QY 241 AASYEGGAGNATAAGAAQAGSGSQRTPPGSCRTFRSAPASAGGYSVOLDALHFFSTFEDS 300

Db 238 ASSYEGGAGASAVAGAQAAPGCGGPRTPPGPCRTFRSGPSSGYSVOLDALHFFPAFEDR 297

QY 301 ALTAMMAOKNLSPLPGSLQPVQENSKTSRSHRSDGFEFSPHSHYSDEAS 354
Db 298 ALTAMMAOKNLSPLPGSLQPVQEDNKTSPRSHRSDGFEFSPHSHYSDEAS 351

RESULT 6

US-10-004-717-46

; Sequence 46, Application US/10004717

; Publication No. US20020192665A1

; GENERAL INFORMATION:

; APPLICANT: ZOGHEI, HUDA Y.

; APPLICANT: YANG, QI

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC USE OF AN

; TITLE OF INVENTION: ATONAL ASSOCIATED SEQUENCE FOR DEAFNESS,

; FILE REFERENCE: P01899054

; CURRENT FILING DATE: 2002-08-16

; PRIOR APPLICATION NUMBER: 09/585,645

; PRIOR FILING DATE: 2000-06-01

; PRIOR APPLICATION NUMBER: 60/176,993

; PRIOR FILING DATE: 2000-01-19

; PRIOR APPLICATION NUMBER: 60/137,060

; PRIOR FILING DATE: 1999-06-01

; NUMBER OF SEQ ID NOS: 69

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 46

; LENGTH: 351

; TYPE: PRT

; ORGANISM: Mus musculus

US-10-004-717-46

Query Match

Best Local Similarity 87.4%; Score 1627.5; DB 13; Length 351;

Matches 313; Conservative 11; Mismatches 27; Indels 3; Gaps 2;

QY 1 MSRLHAEWEAEVKELGDHHRQPHHLPPPPPPPPATLQAREHPVYPPELSLLDSTD 60

Db 1 MSRLHAEWEAEVKELGDHHRQPHHV--PPLTPPPATLQARDLPVYPPELSLLDSTD 58

QY 61 PRAWLAPTLOGICTARAAQYLLHSPGELGASAAAAPRDEVDGREGELVRRSSGASSKSPG 120

Db 59 PRAWLPTTLOGICTARAAQYLLHSPGELGASAAAAPRDEADSQGELVRR--SGCGGLSKSPG 117

QY 121 PVKREOLCKLGGVVDELGCSRQAPSSKQVNGVQKORRLAANARERRRMHGLNHAFF 180

Db 118 PVKREOLCKLGGVVDELGCSRQAPSSKQVNGVQKORRLAANARERRRMHGLNHAFF 177

QY 181 QLRNVIPSFNNDKKLSKYETLQMAQIYINALSELLQTPSGGEGQPPPPASCKSDHHHLRT 240

Db 178 QLRNVIPSFNNDKKLSKYETLQMAQIYINALSELLQTPNVGEGQPPPTASCKNDHHHLRT 237

QY 241 AASYEGGAGNATAAGAAQAGSGSQRTPPGSCRTFRSAPASAGGYSVOLDALHFFSTFEDS 300

Db 238 ASSYEGGAGASAVAGAQAAPGCGGPRTPPGPCRTFRSGPSSGYSVOLDALHFFPAFEDR 297

QY 301 ALTAMMAOKNLSPLPGSLQPVQENSKTSRSHRSDGFEFSPHSHYSDEAS 354

Db 298 ALTAMMAOKNLSPLPGSLQPVQEDNKTSPRSHRSDGFEFSPHSHYSDEAS 351

RESULT 7

US-10-373-249-1

; Sequence 1, Application US/10373249

; Publication No. US20040166091A1

; GENERAL INFORMATION:

; APPLICANT: BROUGH, DOUGLAS E

; TITLE OF INVENTION: MATERIALS AND METHODS FOR TREATING DISORDERS OF THE EAR

; FILE REFERENCE: 214680

; CURRENT APPLICATION NUMBER: US/10/373,249

; CURRENT FILING DATE: 2003-02-24

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-373-249-1

Query Match 87.4%; Score 1627.5; DB 16; Length 351;
Best Local Similarity 88.4%; Pred. No. 2.1e-117;
Matches 313; Conservative 11; Mismatches 27; Indels 3; Gaps 2;

QY 1 MSRLHAEEWAEVKELGDDHRRQPHHLPPPPPPPPATLOAREHPVYPPELSLLDSTD 60
Db 1 MSRLHAEEWAEVKELGDDHRRQPHHV--PPLTPPPATLQARDLPVYPPELSLLDSTD 58
QY 61 PRAWLPTLOGICTARAAQYLLHSPGELGASEAAAPRDEVDGRGELVRRSSGGASSKSPG 120
Db 59 PRAWLTPTLOGICTARAAQYLLHSPGELGASEAAAPRDEADSQGELVRR--SCGGLSKSPG 117
QY 121 PVKREQLCKLGGVVVDDELGCSRQAPSSKQVNGVQKORLAANARERRRMHGLNHAFD 180
Db 118 PVKREQLCKLGGVVVDDELGCSRQAPSSKQVNGVQKORLAANARERRRMHGLNHAFD 177
QY 181 QLRNVIPSFNNDDKLSKYETLQMAQIYINALSELLQTPSGEGEQPPPPASCKSDHHLRT 240
Db 178 QLRNVIPSFNNDDKLSKYETLQMAQIYINALSELLQTPNVGEQPPPTASCNDHHLRT 237
QY 241 AASVEGAGNATAAGAAQAGSGSORPTPGSCRTFRFAPASAGYSVQLDALHFTFEDS 300
Db 238 ASSVEGAGASAVAGAPAGGGPRTPPGPCRTFRFGPSSGGYSVQLDALHFPFEDR 297
QY 301 ALTAMMAQKNSLPSLPGSILOPVEENSKTSPRSHRSDGFEPSHSHYSDSDEAS 354
Db 298 ALTAMMAQKOLSPSLPGSILOPVEDNSKTSPRSHRSDGFEPSHSHYSDSDEAS 351

RESULT 8

US-10-373-249-2

; Sequence 2, Application US/10373249
; Publication No. US20040166091A1
; GENERAL INFORMATION:

; APPLICANT: BROUGH, DOUGLAS E
; TITLE OF INVENTION: MATERIALS AND METHODS FOR TREATING DISORDERS OF THE EAR
; FILE REFERENCE: 214680
; CURRENT APPLICATION NUMBER: US/10/373,249
; CURRENT FILING DATE: 2003-02-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-373-249-2

Query Match 87.4%; Score 1627.5; DB 16; Length 351;
Best Local Similarity 88.4%; Pred. No. 2.1e-117;
Matches 313; Conservative 11; Mismatches 27; Indels 3; Gaps 2;

QY 1 MSRLHAEEWAEVKELGDDHRRQPHHLPPPPPPPPATLOAREHPVYPPELSLLDSTD 60
Db 1 MSRLHAEEWAEVKELGDDHRRQPHHV--PPLTPPPATLQARDLPVYPPELSLLDSTD 58
QY 61 PRAWLPTLOGICTARAAQYLLHSPGELGASEAAAPRDEVDGRGELVRRSSGGASSKSPG 120
Db 59 PRAWLTPTLOGICTARAAQYLLHSPGELGASEAAAPRDEADSQGELVRR--SCGGLSKSPG 117
QY 121 PVKREQLCKLGGVVVDDELGCSRQAPSSKQVNGVQKORLAANARERRRMHGLNHAFD 180
Db 118 PVKREQLCKLGGVVVDDELGCSRQAPSSKQVNGVQKORLAANARERRRMHGLNHAFD 177
QY 181 QLRNVIPSFNNDDKLSKYETLQMAQIYINALSELLQTPSGEGEQPPPPASCKSDHHLRT 240
Db 178 QLRNVIPSFNNDDKLSKYETLQMAQIYINALSELLQTPNVGEQPPPTASCNDHHLRT 237

QY 241 AASVEGAGNATAAGAAQAGSGSORPTPGSCRTFRFAPASAGYSVOLDALHFTFEDS 300
Db 238 ASSVEGAGASAVAGAPAGGGPRTPPGPCRTFRFGPSSGGYSVOLDALHFPFEDR 297
QY 301 ALTAMMAQKNSLPSLPGSILOPVEENSKTSPRSHRSDGFEPSHSHYSDSDEAS 354
Db 298 ALTAMMAQKOLSPSLPGSILOPVEDNSKTSPRSHRSDGFEPSHSHYSDSDEAS 351

RESULT 9

US-10-004-717-40

; Sequence 40, Application US/10004717
; Publication No. US20020192665A1
; GENERAL INFORMATION:

; APPLICANT: ZOGHEI, HUDA Y.
; APPLICANT: YANG, QI
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC USE OF AN
; TITLE OF INVENTION: ATONAL ASSOCIATED SEQUENCE FOR DEAFNESS,
; TITLE OF INVENTION: OSTEOARTHRITIS, AND ABNORMAL CELL PROLIFERATION
; FILE REFERENCE: P01899US4
; CURRENT APPLICATION NUMBER: US/10/004,717
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 09/585,645
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: 60/176,993
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: 60/137,060
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 40
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Zebra Fish
US-10-004-717-40

Query Match 31.3%; Score 582; DB 13; Length 266;
Best Local Similarity 46.5%; Pred. No. 8.1e-37;
Matches 155; Conservative 21; Mismatches 69; Indels 88; Gaps 14;

QY 12 EVKELGDHRRQPHHLPPPPPPPPATLOAREHPVYPPELSLLDSTDPRAMLAFTLOG 71
Db 10 EVVELDVQH-----SSLGREQSKYPPALALMASSDPRAMLAFTLOG 51
QY 72 ICTARAAQYLLHSPGELGASEAAAPRDEVDGRGELVRRSSGGASSKSPGVKREQLCKL 131
Db 52 TCAAH-AYLLHSP--GSS-----AEGVSSASNFRKSSKS--PVKRE-LCRL 93
QY 132 KGGVVVDDELGCSRQAPSSKQVNGVQKORLAANARERRRMHGLNHAFDOLNRVIPSFNN 191
Db 94 KGAVGADE--GRQAPSSKSTNVQKORMAANARERRRMHGLNHAFDLSRVSIPAFDN 150
QY 192 DKLSKYETLQMAQIYINALSELLQTPSGEGEQPPPPASCKSDHHH-LRTAASVEGAGN 250
Db 151 DKLSKYETLQMAQIYINALSDLLQGGAKADPP-----NCDLLHANVLETRDPSRGS-- 203
QY 251 ATAAGAAQAGSGSORPTPGSCRTFRFAPASAGYSVOLDALHFTFEDSALTAMMAQKN 310
Db 204 -----PVCRR-----GTGVGYPYQ-----YEDGTFNSFMEQDL 232
QY 311 LSP---SLPGSILOPVEENSKTSPRSHRSDGE 340
Db 233 QSPSGTSKSGS-----EASKDSPRNSRSDGE 258

RESULT 10

US-10-004-717-66

; Sequence 66, Application US/10004717
; Publication No. US20020192665A1
; GENERAL INFORMATION:

; APPLICANT: ZOGHEI, HUDA Y.
; APPLICANT: YANG, QI
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC USE OF AN

```
; TITLE OF INVENTION: ATONAL ASSOCIATED SEQUENCE FOR DEAFNESS,
; TITLE OF INVENTION: OSTEOARTHRITIS, AND ABNORMAL CELL PROLIFERATION
; FILE REFERENCE: P01899084
; CURRENT APPLICATION NUMBER: US/10/004,717
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 09/585,645
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: 60/176,993
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: 60/137,060
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 66
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Frog
US-10-004-717-66

Query Match      28.8%; Score 537; DB 13; Length 259;
Best Local Similarity 41.8%; Pred. No. 2.3e-33;
Matches 154; Conservative 19; Mismatches 67; Indels 128; Gaps 16;

QY 1 MSRLH-----AEWAUVKELGDHHRQPOPHLPQPPPPPPATLQAREHPVYPPELSLL 56
Db 1 MARLLHGAATAADWCLEL-----PSE--AGLLARDY-----LL 33

QY 57 DSTDPRLMPT-LQICITARAQYLLHSPGLGASEAAAPRDVGRGELVRRSGGASS 115
Db 34 DSSDPRLMPTSLQS-----RPEVYLHPP-----GRAHKVR----- 65

QY 116 SKSPGVKREOLCKLKGVVVDE-----LGCSPQAPSSKQVNGVQKQRL 162
Db 66 -----ELCKLKLGRDDDDDEDEDEERSEGLCHRRPGPGKPGVQKQRL 114

QY 163 ANARERRRMHGLNHAFOQLRVNIPSFNNNDKLSKYETLQMAQIYINALSELLQTPSGGE 222
Db 115 ANARERRRMHGLNHAFOQLRVNIPSFNNNDKLSKYETLQMAQIYINALSELLQTPSGGE 174

QY 223 QPPPPASCKSDHHLRTAAVEGGAGNATAAGAAQASGGSGORPPPGSCRTFRFAPASA 282
Db 175 DPECPPT-----YQLHSGPEPRLVQSGSC-MRFS----- 202

QY 283 GYSVQDLALHSTFEDSALTAMAKNLSPSLPGSILQVQENSKTSPRSHRSDGERS 342
Db 203 GDFPGQ-SPLSTQFQGAALSG-----KGI-----GSAPSSSGEDSKTSPRSHRSDGEP- 251

QY 343 PSHVSDS 350
Db 252 -RSPVSES 258

RESULT 11
US-10-004-717-60
; Sequence 60, Application US/10004717
; Publication No. US20020192665A1
; GENERAL INFORMATION:
; APPLICANT: YANG, QI
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC USE OF AN
; TITLE OF INVENTION: ATONAL ASSOCIATED SEQUENCE FOR DEAFNESS.
; TITLE OF INVENTION: OSTEOARTHRITIS, AND ABNORMAL CELL PROLIFERATION
; FILE REFERENCE: P01899084
; CURRENT APPLICATION NUMBER: US/10/004,717
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 09/585,645
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: 60/176,993
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: 60/137,060
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 60
; LENGTH: 161
; TYPE: PRT
; ORGANISM: chicken
US-10-004-717-60

Query Match      18.1%; Score 337; DB 13; Length 161;
Best Local Similarity 61.2%; Pred. No. 3.6e-18;
Matches 79; Conservative 10; Mismatches 22; Indels 18; Gaps 5;

QY 107 RRSGGAS---SSKSPGVKREOLCKLKGVVVDELGCSPQAPSSKQVNGVQKQRLA 163
Db 42 RPVAGGARVPAGAPGSE---RGAAGARGGG-----GGAGPRA---QVSGVQKQRLA 89

QY 164 ANARERRRMHGLNHAFOQLRVNIPSFNNNDKLSKYETLQMAQIYINALSELLQTPSGGEQ 223
Db 90 ANARERRRMHGLNHAFOQLRVNIPSFNNNDKLSKYETLQMAQIYINALSELLQTPSGGEQ 146

QY 224 PPPPPASCK 232
Db 147 PPEPPAKAE 155

RESULT 12
US-08-722-570-8
; Sequence 8, Application US/08722570
; Publication No. US20030044887A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; TITLE OF INVENTION: NEUROGENIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/722,570
; FILING DATE: 27-SEP-1996
; CLASSIFICATION: 5365
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-63902/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-722-570-8

Query Match      15.1%; Score 281; DB 8; Length 57;
Best Local Similarity 96.5%; Pred. No. 2e-14;
Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 158 KQRLAANARERRRMHGLNHAFOQLRVNIPSFNNNDKLSKYETLQMAQIYINALSELL 214
Db 1 KNRLAANARERRRMHGLNHAFOQLRVNIPSFNNNDKLSKYETLQMAQIYINALSEI 57
```

RESULT 13

US-10-425-259-8
; Sequence 8, Application US/10425259
; Publication No. US20030224431A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; Ma, Qiufu
; TITLE OF INVENTION: NEUROGENIN
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr Hobbach Test Albritton & Herbert LLP
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/425,259
; FILING DATE: 29-Apr-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/932,411A
; FILING DATE: 15-SEP-1997
; APPLICATION NUMBER: US 08/772,009
; FILING DATE: 19-DEC-1996
; APPLICATION NUMBER: US 08/722,570
; FILING DATE: 19-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-63902-3/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-425-259-8

Query Match 15.1%; Score 281; DB 12; Length 57;
Best Local Similarity 96.5%; Pred. No. 2e-14; Indels 0; Gaps 0;
Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 158 KORLAANARRRRMGLHAFDQLRNVPSPFNNKKLSKYETLQMAQIYNALSEL 214
DB 1 KNRLLAANARRRRMGLHAFDQLRNVPSPFNNKKLSKYETLQMAQIYNALSEL 57

RESULT 14

US-10-004-717-19
; Sequence 19, Application US/10004717
; Publication No. US20020192665A1
; GENERAL INFORMATION:
; APPLICANT: ZOGHBI, HUDA Y.
; APPLICANT: YANG, QI
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC USE OF AN
; TITLE OF INVENTION: ATONAL ASSOCIATED SEQUENCE FOR DEAFNESS,
; FILE REFERENCE: P01899US4
; CURRENT APPLICATION NUMBER: US/10/004,717
; CURRENT FILING DATE: 2002-08-16

; PRIOR APPLICATION NUMBER: 09/585,645
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: 60/176,993
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: 60/137,060
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 19
; LENGTH: 178
; TYPE: PRT
; ORGANISM: chicken
US-10-004-717-19

Query Match 12.3%; Score 228.5; DB 13; Length 178;
Best Local Similarity 34.7%; Pred. No. 9.7e-10;
Matches 77; Conservative 26; Mismatches 54; Indels 65; Gaps 11;
QY 109 SSGGASSKSPGPVKVREQLCKLKGVVVDLGCRRQRAPSSKQVGVQKORLLAANARE 168
DB 7 SSGGV--SEPPGAPRRR---RRRG-----RARARTEALLHTLKSRVRKANDRE 51
QY 169 RRRMGLHAFDQLRNVPSPFNNKKLSKYETLQMAQIYNALSELQTPSGGSEQ--PPP 226
DB 52 RNRMHHLNAALDELRSVLPTFPDDTKLTKTETLRFAYNIWALSETLRL---AQQLPPP 108
QY 227 P-----PASCSDHHLRTAASYEGGAGNATAAGAAQQAQSGGSGRTPPGSCHTRESAP 279
DB 109 PATRGPPAPSPGSD-----AGSWLSSGS---PAAPSLC----- 139
QY 280 ASAGGYSVQLDALHFTFEDSALT---AMMAQKNLSPLSPGS 318
DB 140 ASAGSPSP-----ATSEDCGVPSDALRAFRGLPAAFGA 175

RESULT 15

US-10-004-717-21
; Sequence 21, Application US/10004717
; Publication No. US20020192665A1
; GENERAL INFORMATION:
; APPLICANT: ZOGHBI, HUDA Y.
; APPLICANT: YANG, QI
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC USE OF AN
; TITLE OF INVENTION: ATONAL ASSOCIATED SEQUENCE FOR DEAFNESS,
; FILE REFERENCE: P01899US4
; CURRENT APPLICATION NUMBER: US/10/004,717
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 09/585,645
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: 60/176,993
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: 60/137,060
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 21
; LENGTH: 213
; TYPE: PRT
; ORGANISM: chicken
US-10-004-717-21

Query Match 12.1%; Score 224.5; DB 13; Length 213;
Best Local Similarity 29.7%; Pred. No. 2.5e-09;
Matches 78; Conservative 24; Mismatches 88; Indels 73; Gaps 9;
QY 32 PPPQPATLQAREHPVYPPELSLDSTDPRAWLFTLQICTARAAQYLLHSPGLGASE 91
DB 2 PVKARSPA-----PAAEDELLRLASP---AP-----SASLPSSA 34
QY 92 AAAPRDEVDGRGELVRRSSGSSKSPGVKVRQELCKLKGVVVDLGCRRQRAPSSK 151
DB 35 GDEDEDEDGRPR--RLOEGARRAGRGQRPRA-----ARTAE 70

Qy	152	QVNGVQKORRLAANAHERRMHCLNHAFDQLRNVI	PSFNDKLSKYETLOMAQIYINAL	211
Db	71	TAQRIKRSRRLKANNRERNMNLNAALDALRDVLP	TFPEDAKLTKTETLRFHNYIWAL	130
Qy	212	SELLQTPS----	GGEQPPPPASCKSDHHLRTAASVEGGAGNATAAG-----A	256
Db	131	TETLRLAGARLGGADAAPGAAAEKSP----	SPASSWSGGASPASPACTLSFGSPA	187
Qy	257	QOASGGSQRPTPPGSCRTFSAP		279
Db	188	GSASDAEHWPFRG----	RFAPP	206

Search completed: September 21, 2004, 21:40:26
Job time : 163.424 secs

Blank
Sheet